

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2003, 05:56:31 / Search time 8288 Seconds
(without alignments)

11219.544 Million cell updates/sec

Title: US-09-938-803-26

Perfect score: 2273

Sequence: 1 59999999gaagcgatagctt.....ttttcaaaaaaaaaa 2273

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sfr:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sfr:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_rin:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2273	100.0	2273	6	AX048067 Sequence
2	2211	97.3	2264	6	AX073570 Sequence
3	1996	87.8	2058	6	BC000664 Homo sapi
4	1868	82.2	2482	6	AK096775 Homo sapi
5	1842	81.0	1842	6	AX331926 Homo sapi
6	1842	81.0	1842	6	AX410769 Sequence
7	1842	81.0	1842	6	HSU72515
8	1461	64.3	1461	6	AX073560 Sequence
9	1258	54.5	1573	9	AK058063 Homo sapi
10	1094	48.1	1146	9	BT007000 Homo sapi
11	1093	48.1	1146	12	BT007735 Synthetic
12	708	31.1	155975	9	AC006512 Homo sapi
13	708	31.1	222930	9	HSU47924
14	424	18.7	1641	9	BC009856 Homo sapi
15	402	17.7	450	6	BD023986 Sequence
16	254	11.2	990	6	AX331531 Sequence
17	254	11.2	990	6	AX331927 Sequence
18	254	11.2	990	6	AX335065 Sequence
19	254	11.2	990	6	HSU72517
20	112	4.9	1414	6	AX336765
21	112	4.9	1414	6	HSU72507
22	73	3.2	280	6	AX409345 Sequence
23	64	2.8	120846	9	AL365232 Human DNA
24	60	2.6	37223	9	AF135116 Homo sapi
25	60	2.6	74245	2	AC022622
26	60	2.6	101805	9	AC073508 Homo sapi
27	60	2.6	107532	9	AC010425 Homo sapi
28	60	2.6	122900	2	AL139126 Homo sapi
29	60	2.6	125527	9	AL353133 Human DNA
30	60	2.6	132557	9	AC004520 Homo sapi
31	60	2.6	137830	9	AC003029
32	60	2.6	141923	2	AC109997 Homo sapi
33	60	2.6	145887	9	AC091435 Homo sapi
34	60	2.6	149089	9	AC090735 Homo sapi
35	60	2.6	154381	2	AC009096 Homo sapi
36	60	2.6	156300	9	AC005919 Homo sapi
37	60	2.6	158832	2	AC015475 Homo sapi
38	60	2.6	160597	9	AL590666 Human DNA
39	60	2.6	168137	9	AC090283 Homo sapi
40	60	2.6	175088	9	AC121766 Homo sapi
41	60	2.6	176845	9	AC040163 Homo sapi
42	60	2.6	181188	2	AC067983 Homo sapi
43	60	2.6	184133	9	AC103559 Homo sapi
44	60	2.6	194385	9	CNS00001
45	60	2.6	200807	9	AC073278 Homo sapi

ALIGNMENTS

RESULT 1	AX048067	2273 bp	DNA	linear	PAT 15-DEC-2000
LOCUS	AX048067				
DEFINITION	Sequence 61 from Patent WO0070047.				
ACCESSION	AX048067				
VERSION	AX048067.1				
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Yue, H., Tang, Y.T., Lal, P., Reddy, R., Batra, S., Baughn, M.R., Yang, J., Azimzai, Y., Lu, D.A., Au-Young, J., and Shih, L.L.				
TITLE	Full-length molecules expressed in human tissues				

JOURNAL Patent: WO 0070047-A 61 23-NOV-2000;
Incyte Genomics, Inc. (US)
Location/Qualifiers
source 1..2273
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1928920CB1"
BASE COUNT 494 a 630 c 547 g 602 t
ORIGIN
Query Match 100.0%; Score 2273; DB 6; Length 2273;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGGTGAGACGATACGTTTGGCCCGCATTCGGGGCGCGGACCTGGGGGGGTCCCTGTG 60
Db 1 GGGGGTGAGACGATACGTTTGGCCCGCATTCGGGGCGCGGACCTGGGGGGGTCCCTGTG 60
QY 61 GGGCTCCCGAGTTAAGATGGCGTCTCAGCGAGGGGGAGCAGGGGACTGTGTGGCGC 120
Db 61 GGGCTCCCGAGTTAAGATGGCGTCTCAGCGAGGGGGAGCAGGGGACTGTGTGTGGCGC 120
QY 121 TGGGGGGGGTTTGCAGTGGGGTTTCCAGAGCTGAGCTTAAACAAGTTGGCGACGTGCC 180
Db 121 TGGGGGGGGTTTGCAGTGGGGTTTCCAGAGCTGAGCTTAAACAAGTTGGCGACGTGCC 180
QY 181 TGGGGCGGTGAGAACAGGCGCTGGGCTGATCATCTTCATCTTCTGGGTTACCCCTTGG 240
Db 181 TGGGGCGGTGAGAACAGGCGCTGGGCTGATCATCTTCATCTTCTGGGTTACCCCTTGG 240
QY 241 CTTTGTGTTATGGGATTAACCTTTTTCACAGAGACCTTACCTATCACCCTCTTCATA 300
Db 241 CTTTGTGTTATGGGATTAACCTTTTTCACAGAGACCTTACCTATCACCCTCTTCATA 300
QY 301 CCTTACAGGCTCTCAATGCTTATTTTAACTTTGAAACCAAGCTTACCACTCCCTGCG 360
Db 301 CCTTACAGGCTCTCAATGCTTATTTTAACTTTGAAACCAAGCTTACCACTCCCTGCG 360
QY 361 TGTGTATTTGTCTCAGTCTCTCATCTTGCATTAATGGGCGGACCATCATCGCGCTCC 420
Db 361 TGTGTATTTGTCTCAGTCTCTCATCTTGCATTAATGGGCGGACCATCATCGCGCTCC 420
QY 421 TCACTACCTTTTGTCTTCCAGATGGCTTACCTTCTGGCTGATATTAATCACTGCCAC 480
Db 421 TCACTACCTTTTGTCTTCCAGATGGCTTACCTTCTGGCTGATATTAATCACTGCCAC 480
QY 481 GCAACTACGATATCAATGAGCAATGCACTTGTCTGACTTTGAAGCTGATTGGTT 540
Db 481 GCAACTACGATATCAATGAGCAATGCACTTGTGTCTGACTTTGAAGCTGATTGGTT 540
QY 541 TGGCTGTGACTATCTTTGACGAGGAGAAAGATCGAATTCCTTGTCTCTGAGCAACAGA 600
Db 541 TGGCTGTGACTATCTTTGACGAGGAGAAAGATCGAATTCCTTGTCTCTGAGCAACAGA 600
QY 601 AATATGCGCATACGTTGTCTTCTCTGCTGGAAGTTGTGTGTTCTCTACTTCTATG 660
Db 601 AATATGCGCATACGTTGTCTTCTCTGCTGGAAGTTGTGTGTTCTCTACTTCTATG 660
QY 661 GGGGCTTCTTGTGTGGGCGCCAGTTCTCAATGAATCACTACATAGAGCTGTGAGGAG 720
Db 661 GGGGCTTCTTGTGTGGGCGCCAGTTCTCAATGAATCACTACATAGAGCTGTGAGGAG 720
QY 721 AGCTGATTTGACATACAGAAAGATACCAACAGCATCATTCCTGCTCTCAAGCGCTGA 780
Db 721 AGCTGATTTGACATACAGAAAGATACCAACAGCATCATTCCTGCTCTCAAGCGCTGA 780
QY 781 GTCTGGGCTTTTCTACTAGTGGCTACACAGCTCAAGCCCAATACACAGAAACT 840
Db 781 GTCTGGGCTTTTCTACTAGTGGCTACACAGCTCAAGCCCAATACACAGAAACT 840
QY 841 ATCTCTCACTGAAGACTATGACAAACCCCTTCTGTTCCGTGATGATCACTGCTGA 900
Db 841 ATCTCTCACTGAAGACTATGACAAACCCCTTCTGTTCCGTGATGATCACTGCTGA 900

Db 841 ATCTCTCACTGAAGACTATGACAAACCCCTTCTGTTCCGTGATGATCACTGCTGA 900
QY 901 TCTGGGGCAAGTTTGTCTGTACAAATATGTCACTGTGGCTGTACAGAAAGATAT 960
Db 901 TCTGGGGCAAGTTTGTCTGTACAAATATGTCACTGTGGCTGTACAGAAAGATAT 960
QY 961 GCATTTTGAACGGGCTGGGCTTCAATGCTTTGAAAGAAAAGGCAAGCAAGTGGAGT 1020
Db 961 GCATTTTGAACGGGCTGGGCTTCAATGCTTTGAAAGAAAAGGCAAGCAAGTGGAGT 1020
QY 1021 CCTGTGCCAAATGAAGTGTGGCTTTGAAACCAACCCCGCTTCACTGAGCACTTG 1080
Db 1021 CCTGTGCCAAATGAAGTGTGGCTTTGAAACCAACCCCGCTTCACTGAGCACTTG 1080
QY 1081 CCTCATTTCAACATCAACCAACGCTGGGTGGCCGCTACATCTTCAACGACTCAAGT 1140
Db 1081 CCTCATTTCAACATCAACCAACGCTGGGTGGCCGCTACATCTTCAACGACTCAAGT 1140
QY 1141 TCTTGGAAATTAAGAACTCTCAGAGTCTCTGTTGCTATTCCTGGGCTCTGGCAG 1200
Db 1141 TCTTGGAAATTAAGAACTCTCAGAGTCTCTGTTGCTATTCCTGGGCTCTGGCAG 1200
QY 1201 GCTGCACTCAGGATACCTGTCTGCTCAGATGGAATTCCTCATTTGTTGGA 1260
Db 1201 GCTGCACTCAGGATACCTGTCTGCTCAGATGGAATTCCTCATTTGTTGGA 1260
QY 1261 GACAGGCTGCAAGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCCGCTATTA 1320
Db 1261 GACAGGCTGCAAGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCCGCTATTA 1320
QY 1321 TCTTCAAGCCCTTCTATATTTGTTGTCACAGACATTCACATGCTTCTCATGAGTTACT 1380
Db 1321 TCTTCAAGCCCTTCTATATTTGTTGTCACAGACATTCACATGCTTCTCATGAGTTACT 1380
QY 1381 CCATGACTGCTTCTGCTCTTCACTGGGCAAAATGCTTAAAGTATTAATTCATCT 1440
Db 1381 CCATGACTGCTTCTGCTCTTCACTGGGCAAAATGCTTAAAGTATTAATTCATCT 1440
QY 1441 ATTTCTTGGCCATCTTCTTCTGAGCTTACTATTTGCTTATTTTCAACAG 1500
Db 1441 ATTTCTTGGCCATCTTCTTCTGAGCTTACTATTTGCTTATTTTCAACAG 1500
QY 1501 CAATGTCGCCAAGAAAGAAAGTAAAGATGAATTAATTCATTTCCCTGCTGCT 1560
Db 1501 CAATGTCGCCAAGAAAGAAAGTAAAGATGAATTAATTCATTTCCCTGCTGCT 1560
QY 1561 GTGCGGGAAGTGTGCAAGAACTACTGCTCTCTTTCACAGCACTCTTTGCCCGCAG 1620
Db 1561 GTGCGGGAAGTGTGCAAGAACTACTGCTCTCTTTCACAGCACTCTTTGCCCGCAG 1620
QY 1621 CAGAGATGGAAGAACCGGAGAGTGAAGATGATGCTTCCAGCTGTGCTGCTGCC 1680
Db 1621 CAGAGATGGAAGAACCGGAGAGTGAAGATGATGATGCTTCCAGCTGTGCTGCTGCC 1680
QY 1681 AGCCAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGGAAGAGGCTGTGCTT 1740
Db 1681 AGCCAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGGAAGAGGCTGTGCTT 1740
QY 1741 GTCAACCAACGCTGGAATGCAAGTGGCGGATCTCAGCTCAACCGAACCCTCCACTCTGG 1800
Db 1741 GTCAACCAACGCTGGAATGCAAGTGGCGGATCTCAGCTCAACCGAACCCTCCACTCTGG 1800
QY 1801 TTCAAGTATTTTCTGCTCAGCTTCCCAAGTACTGGGAATACAGGCAAGCAACATG 1860
Db 1801 TTCAAGTATTTTCTGCTCAGCTTCCCAAGTACTGGGAATACAGGCAAGCAACATG 1860
QY 1861 CCCAGCTAATTTTGTATTTTCAAGTGAAGAGGGAATTCACAGTTGGCAGGCTGCT 1920
Db 1861 CCCAGCTAATTTTGTATTTTCAAGTGAAGAGGGAATTCACAGCTTTGGCAGGCTGCT 1920
QY 1921 TCGAATCTCTGACCGGACGATGATCAACCCGCTTCCGCTTCCAAAGTGTGGATTA 1980
Db 1921 TCGAATCTCTGACCGGACGATGATCAACCCGCTTCCGCTTCCAAAGTGTGGATTA 1980


```

Db      1254 TGTATTGAGAAAGACAGGCTGCCAGGCTCATTCAGAAAGGCCACCTTGAGCAAGCT 1313
Qy      1307 GGGCCGCAATTAAGTCTCTCAGGCTTCTACTATTGGTGCACAGACCATTCACAGGCT 1366
Db      1314 GGGCCGCAATTAAGTCTCTCAGGCTTCTACTATTGGTGCACAGACCATTCACAGGCT 1373
Qy      1367 CTTGATGGGTTACTTCATGACTGCTTGCCTTTTCACGTGGACCAATGGCTTAAGT 1426
Db      1374 CTTGATGGGTTACTTCATGACTGCTTGCCTTTTCACGTGGACCAATGGCTTAAGT 1433
Qy      1427 GTTAAATCCATTAATTTCTTGGGCAACATCTTCCCTGAGCTTCTATTCATTTGCC 1486
Db      1434 GTTAAATCCATTAATTTCTTGGGCAACATCTTCCCTGAGCTTCTATTCATTTGCC 1493
Qy      1487 TTATATTCACAAAGCAATGAGTGCAGAAAGAAAGATTAAAGATGATATTCAT 1546
Db      1494 TTATATTCACAAAGCAATGAGTGCAGAAAGAAAGATTAAAGATGATATTCAT 1553
Qy      1547 TTCCTGTGGCTGTGGGAGCTGTGCAGAAATCTACTGCTCTCCTTTTCACAGCACT 1606
Db      1554 TTCCTGTGGCTGTGGGAGCTGTGCAGAAATCTACTGCTCTCCTTTTCACAGCACT 1613
Qy      1607 CTTTGGCCCAAGAGAGAAATGGAAGAGCAGGAGGTGGAAGATGATGCTTCCAGCT 1666
Db      1614 CTTTGGCCCAAGAGAGAAATGGAAGAGCAGGAGGTGGAAGATGATGCTTCCAGCT 1673
Qy      1667 GTGCTCTGTGCTGCAGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGGGAGA 1726
Db      1674 GTGCTCTGTGCTGCAGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGGGAGA 1733
Qy      1727 AGGCGTCTTGTGTCACCCAGCGTGGAAATGCAATGCGGAGTCTCAGCTCACCGCAAC 1786
Db      1734 AGGCGTCTTGTGTCACCCAGCGTGGAAATGCAATGCGGAGTCTCAGCTCACCGCAAC 1793
Qy      1787 CTCACCTCTGCGGTTCAAGTGAATTTTCGCGCTCAGCCTCCCAAGTATGCGGGAATACA 1846
Db      1794 CTCACCTCTGCGGTTCAAGTGAATTTTCGCGCTCAGCCTCCCAAGTATGCGGGAATACA 1853
Qy      1847 GGCAGCGCACCATGCGCAGCTAATTTTGTATTTTCAGTAAACGGGATTTCCACACGT 1906
Db      1854 GGCAGCGCACCATGCGCAGCTAATTTTGTATTTTCAGTAAACGGGATTTCCACACGT 1913
Qy      1907 TGGCAGGCTGTGCTGCAATCTCTGACCGCAAGTATCCACCGCTTCCGCTCCCAAG 1966
Db      1914 TGGCAGGCTGTGCTGCAATCTCTGACCGCAAGTATCCACCGCTTCCGCTCCCAAG 1973
Qy      1967 TGGTGGGATTACAGGCGTGAAGCAACGTCGCGCGGCCCAAAAGGGGAACTCTTGAGGAGG 2026
Db      1974 TGGTGGGATTACAGGCGTGAAGCAACGTCGCGCGGCCCAAAAGGGGAACTCTTGAGGAGG 2033
Qy      2027 AGCAGAGGGGCTCATCTCTCCCTCTGATTTCCCATGACATTTGCTTATCTCTCCCA 2086
Db      2034 AGCAGAGGGGCTCATCTCTCCCTCTGATTTCCCATGACATTTGCTTATCTCTCCCA 2093
Qy      2087 TCTAGCCAGGAATTAATTTGTTTTTCTTCTGCGCAATTTACTATGATTTGTAATGTCGG 2146
Db      2094 TCTAGCCAGGAATTAATTTGTTTTTCTTCTGCGCAATTTACTATGATTTGTAATGTCGG 2153
Qy      2147 CTACCCACACCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTCCCACTT 2206
Db      2154 CTACCCACACCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTCCCACTT 2213
Qy      2207 TTCTTACCTTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2257
Db      2214 TTCTTACCTTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2264

```

RESULT 3
 BC000664
 LOCUS
 DEFINITION Homo sapiens putative protein similar to nesy (Drosophila), mRNA
 (CDNA clone MGC:1311 IMAGE:3349388), complete cds.

ACCESSION BC000664
 VERSION BC000664.1 GI:12653756
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 1 (bases 1 to 2058)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Mang, J., Heich, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Useth, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL MEDLINE
 PUBMED 22388257
 REFERENCE
 2 (bases 1 to 2058)
 Strausberg, R.
 Direct Submision
 Submitted (15-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Khom, P., Latic, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskell, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

FEATURES
 source
 1. 2058
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:1311 IMAGE:3349388"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NIH MGC 15"
 /lab_host="DH10B-R"
 /note="Vector: POTB7"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 5 Row: d Column: 14
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15375317.

gene 1. .2058
/gene="C3F"
/db_xref="LocusID:10162"
177. .1322
/codon_start=1
/product="putative protein similar to nesey"
/protein_id="AAH0664.1"
/db_xref="GI:12653757"
/db_xref="LocusID:10162"
/translation="MGRTITAVLTFCFQMAVLLAGYYTATGNVDIKMTMPHCULTL
KILGLADVDFDGGKDNLSSEBOOKYAIIGVPSLLEVAESFYTGKFLVGPQSNMNY
MLVQSELIDIPKIPNSITIPALRSLSLFVLVGTLLSPHITLDELDNHP
WRCKMILMGKFLVLYKVTCLWLTGVCILGLGNGEKEKAKMDACANNKWLIF
ETNPRFTGTAJENINTAMVARYIFKRLFKLENKELSGSLFLFALMGLHSGYLV
CEQMFLIIVEROARLLOESPTLSKLAITVLQPFYLVQOTIHLFMGYSMTAFC
LFTMDKMLKYYSIFLGHIFLSLFIPIYIKHAWPKKKKKME"

BASE COUNT 470 a 575 c 457 g 556 t

ORIGIN

Query Match 87.8%; Score 1996; DB 9; Length 2058;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 227 GGGTTACCCCTTGGCTTTGTTTATGGCACTTCTTTCTACAGAGACTCACTCAT 286
DB 8 GGGTTACCCCTTGGCTTTGTTTATGGCACTTCTTTCTACAGAGACTCACTCAT 67

QY 287 CCACCTCTTCACATACCTTTACAGGGCTCTCAATGCTTATTTTAACTTTGGAAACAGCT 346
DB 68 CCACCTCTTCACATACCTTTACAGGGCTCTCAATGCTTATTTTAACTTTGGAAACAGCT 127

QY 347 CTACCACTCCCTGCTGTGATGCTTGAAGTTCTCATCTTGAATGAGGCGCGAC 406
DB 128 CTACCACTCCCTGCTGTGATGCTTGAAGTTCTCATCTTGAATGAGGCGCGAC 187

QY 407 CATCACTGCGCTCTCACTACCTTTGCTTCCAGATGCGCTTCTGCTGATATCTA 466
DB 188 CATCACTGCGCTCTCACTACCTTTGCTTCCAGATGCGCTTCTGCTGATATCTA 247

QY 467 TTACACTGCGCGCACTGATATCAAGTGAAGTGAATGCAATGCTTCTGCTT 526
DB 248 TTACACTGCGCGCACTGATATCAAGTGAAGTGAATGCAATGCTTCTGCTT 307

QY 527 GAAGCTGATGTTGGCTGTGATCTACTTTCAGAGGAGAAAGATCAGAAATCTCTGTC 586
DB 308 GAAGCTGATGTTGGCTGTGATCTACTTTCAGAGGAGAAAGATCAGAAATCTCTGTC 367

QY 587 CTCTGAGCAAGAAATATGCCATACGTGCTGTTCTTCCCTGCTGGAAGTGTGCTT 646
DB 368 CTCTGAGCAAGAAATATGCCATACGTGCTGTTCTTCCCTGCTGGAAGTGTGCTT 427

QY 647 CTCTACTCTTATGAGGCGCTTCTGCTAGAGGCGCGCACTTCAATGAATACATCACTGA 706
DB 428 CTCTACTCTTATGAGGCGCTTCTGCTAGAGGCGCGCACTTCAATGAATACATCACTGA 487

QY 707 GCTGTGCAAGGAGAGCTGATGATACATACAGAGAAAGATACCAACAGCATCTCTGTC 766
DB 488 GCTGTGCAAGGAGAGCTGATGATACATACAGAGAAAGATACCAACAGCATCTCTGTC 547

QY 767 TCTCAAGCGCTGATCTGAGGCGCTTTCTAAGTGGCTACACCTGCTGAGCGCCCA 826
DB 548 TCTCAAGCGCTGATCTGAGGCGCTTTCTAAGTGGCTACACCTGCTGAGCGCCCA 607

QY 827 CATCAGAGAGACTATCTCTCAGTGAAGTATGACACCAACCCCTTGTGTTCCGCTG 886
DB 608 CATCAGAGAGACTATCTCTCAGTGAAGTATGACACCAACCCCTTGTGTTCCGCTG 667

QY 887 CATGTAACATGCTGATCTGGGCAAGTTTGTGCTGTACAAATATGTCACTGTTGGCTGT 946
DB 668 CATGTAACATGCTGATCTGGGCAAGTTTGTGCTGTACAAATATGTCACTGTTGGCTGT 727

QY 947 CACAGAGAGATGATGATTTTGAAGGCGCTGAGCTTCAATGCTTTGAAGAAAGGCGCA 1006

DB 728 CACAGAGAGATGATGATTTTGAAGGCGCTGAGCTTCAATGCTTTGAAGAAAGGCGCA 787

QY 1007 GGCAGAGTGGATGCTTGTGCGCAACATGAAGTGTGCTTTTGAAGAAACCCCGCTT 1066

DB 788 GGCAGAGTGGATGCTTGTGCGCAACATGAAGTGTGCTTTTGAAGAAACCCCGCTT 847

QY 1067 CACTGCAACATGCTCTCATTTCAACATCAACCAAGCGCTGGGTGGCGCTACATCTT 1126

DB 848 CACTGCAACATGCTCTCATTTCAACATCAACCAAGCGCTGGGTGGCGCTACATCTT 907

QY 1127 CAAAGACTCAAGTCTCTTGAAGAAATGAAGAACTCTCAGAGTCTCTGTTGATTTCT 1186

DB 908 CAAAGACTCAAGTCTCTTGAAGAAATGAAGAACTCTCAGAGTCTCTGTTGATTTCT 967

QY 1187 GGCCTCTGCGACGCTCTCAGATGATGCTGCTGCTTCCATGGAATTTCTCAT 1246

DB 968 GGCCTCTGCGACGCTCTCAGATGATGCTGCTGCTTCCATGGAATTTCTCAT 1027

QY 1247 TGTATTGTGAAAGACAGGCTGCGAGCTCATTTCAAGAGAGCCGCTGAGCAAGCT 1306

DB 1028 TGTATTGTGAAAGACAGGCTGCGAGCTCATTTCAAGAGAGCCGCTGAGCAAGCT 1087

QY 1307 GGCCTGATTAAGTCTCTCAGGCTTCTAATTTGTGCAACAGACCATCCAGTGGCT 1366

DB 1088 GGCCTGATTAAGTCTCTCAGGCTTCTAATTTGTGCAACAGACCATCCAGTGGCT 1147

QY 1367 CTTCATGAGTTACTCCATGATGCTCTTCTGCTCTTCAAGTGGACAAATGCTTAAGT 1426

DB 1148 CTTCATGAGTTACTCCATGATGCTCTTCTGCTCTTCAAGTGGACAAATGCTTAAGT 1207

QY 1427 GTATAAATCATATTTCTTGGGCAATCTTCTGAGGCTCTATTCATATTGGC 1486

DB 1208 GTATAAATCATATTTCTTGGGCAATCTTCTGAGGCTCTATTCATATTGGC 1267

QY 1487 TTATATTCAAAAGCAATGCTGCCAAGAAAGAAAGATTAAAGATGAATTAATCAT 1546

DB 1268 TTATATTCAAAAGCAATGCTGCCAAGAAAGAAAGATTAAAGATGAATTAATCAT 1327

QY 1547 TTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1606

DB 1328 TTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1387

QY 1607 CCTTTGCCCAAGAGAGAAATGAGAAAGCAAGGAGTGAAGATCGATGCTTCAGCT 1666

DB 1388 CCTTTGCCCAAGAGAGAAATGAGAAAGCAAGGAGTGAAGATCGATGCTTCAGCT 1447

QY 1667 GTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1726

DB 1448 GTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1507

QY 1727 AGGCGCTTGTGCTTGTCAACCCAGCTGGAATGCAATGCTGCTGCTGCTGCTGCT 1786

DB 1508 AGGCGCTTGTGCTTGTCAACCCAGCTGGAATGCAATGCTGCTGCTGCTGCTGCTGCT 1567

QY 1787 CTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1846

DB 1568 CTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1627

QY 1847 GGCAGCGCAACATGCGCCAGCTAATTTTGTATTTTCAATGAAGAGGAAATTCACACGT 1906

DB 1628 GGCAGCGCAACATGCGCCAGCTAATTTTGTATTTTCAATGAAGAGGAAATTCACACGT 1687

QY 1907 TGGCGAGCTGCTGCTGCAATCTCTGACCGCAAGTATCAACCCGCTCGCTCCCAAG 1966

DB 1688 TGGCGAGCTGCTGCTGCAATCTCTGACCGCAAGTATCAACCCGCTCGCTCCCAAG 1747

QY 1967 TGTGGGATTAAGGCTGAGGCAAGGCTGAGGCGGCGGCAAGGAGAAATCTTGTGGAGG 2026

DB 1748 TGTGGGATTAAGGCTGAGGCAAGGCTGAGGCGGCGGCAAGGAGAAATCTTGTGGAGG 1807

QY 2027 AGCAGAGGAGCTCACATCTCCCTCTGATTTCCCGCATGCAATTCCTTATCTCCCA 2086

DB 1808 AGCAGAGGAGCTCACATCTCCCTCTGATTTCCCGCATGCAATTCCTTATCTCCCA 1867

Qy	2087	CTGAGCCAGGAATCTATTGCTGTTTTTCTCTCCCAATTACTATGATGTTGCTAGTCGCCG	2146
Db	1868	CTGAGCCAGGAATCTATTGCTGTTTTTCTCTCCCAATTACTATGATGTTGCTAGTCGCCG	1927
Qy	2147	CTACACCAACCCCCCCCATGGGGGGGTGGAAGGGGTTGCAAGGCCCTGCTCCACTT	2206
Db	1928	CTACACCAACCCCCCCCATGGGGGGGTGGAAGGGGTTGCAAGGCCCTGCTCCACTT	1987
Qy	2207	TTTCTACCTTGGAACTGATTAGATTAATAATCACTTCTGTTTGTTCAGTTTTCAAAAAA	2266
Db	1988	TTTCTACCTTGGAACTGATTAGATTAATAATCACTTCTGTTTGTTCAGTTTTCAAAAAA	2047
Qy	2267	AAAAAA 2273	
Db	2048	AAAAAA 2054	
RESULT 4			
LOCUS	AK096775	2482 bp	mRNA linear PRI 15-JUL-2002
DEFINITION	Homo sapiens cDNA FLJ39456 fis, clone PROST2010782, highly similar to Human C3f mRNA.		
ACCESSION	AK096775		
VERSION	AK096775.1	GI:21756342	
KEYWORDS	oligo capping, fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Ninomiya, K., Nagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Seto, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Karakami, B., Suzuki, Y., Sugano, S., Nagahata, K., Masuno, Y., Nagai, K. and Isogai, T.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2482)		
AUTHORS	Isogai, T. and Yamamoto, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FIJ Project (HRI Team) ; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genom@cs.hri.co.jp; Tel:81-438-52-3975; Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center; National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.		
FEATURES			
source	location/Qualifiers		
	1..2482		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="PROST2010782"		
	/tissue_type="prostate"		
	/clone_id="PROST2"		
	/note="Cloning vector: pME18SFL3"		
BASE COUNT	526 a 687 c 615 g	654 t	
ORIGIN			
Query Match	82.2%	Score 1668;	DB 9; Length 2482;
Best Local Similarity	99.8%	Pred. No. 0;	
Matches 2208;	Conservative 0;	Mismatches 2;	Indels 2; Gaps 2

Db	273	GGGGGGTCCCTGTTGGGCTCCCGAGTTAAAGATGCGTCTTCAAGCGAAGGGGACAGGG	332
Qy	107	GACTGTGGTGGCGCTGGCGGGGGTTTCGCAAGTCGGGTTTCCAGAGCTGAGCCTTAA	166
Db	333	GACTGTGGTGGCGCTGGCGGGGGTTTCGCAAGTCGGGTTTCCAGAGCTGAGCCTTAA	392
Qy	167	GTTGGCAGCGTCCCTGGGCGCGCTCAGAAACAGCGCTCGGCTGATCATCTTCATCTT	226
Db	393	GTTGGCAGCGTCCCTGGGCGCGCTCAGAAACAGCGCTCGGCTGATCATCTTCATCTT	452
Qy	227	GGGTTACCCCTTGGCTTGTGTTTATGGGCAATTCCTTTGACAAAGAGACTTCCTCAT	286
Db	453	GGGTTACCCCTTGGCTTGTGTTTATGGGCAATTCCTTTGACAAAGAGACTTCCTCAT	512
Qy	287	CCACCTCTTCATACCTTTACAGGCGCTCCTCAATGCTTAATTTAACTTTGAAACAG	346
Db	513	CCACCTCTTCATACCTTTACAGGCGCTCCTCAATGCTTAATTTAACTTTGAAACAG	572
Qy	347	CTACCACTCCCTGCTGTGTATGTGTGCTTCAAGTTCCTCATCTTGCCTAATGGCG	406
Db	573	CTACCACTCCCTGCTGTGTATGTGTGCTTCAAGTTCCTCATCTTGCCTAATGGCG	632
Qy	407	CATACATGCCGCTCCTCACTACCTTTGGCTCCCAATGGCCATCTTCGGCTGATACT	466
Db	633	CATACATGCCGCTCCTCACTACCTTTGGCTCCCAATGGCCATCTTCGGCTGATACT	692
Qy	467	TTACACTGCCACCGGCAACTGATATCAAGTGAAGATGCAATGCACTTGTTCGACTT	526
Db	633	TTACACTGCCACCGGCAACTGATATCAAGTGAAGATGCAATGCACTTGTTCGACTT	752
Qy	527	GAAGCTGATGTTGGCTGTTGACTACCTTGTGACGAGGGAAGATCAGAAATTCCTG	586
Db	753	GAAGCTGATGTTGGCTGTTGACTACCTTGTGACGAGGGAAGATCAGAAATTCCTG	812
Qy	587	CTTCGAGGAAGAAATATGCGCAATCGTGGTCCCTCCCTGGTGAAGTGTGGTGT	646
Db	813	CTTCGAGGAAGAAATATGCGCAATCGTGGTCCCTCCCTGGTGAAGTGTGGTGT	872
Qy	647	CTCCACTTCTATGGGGCTTCTTGATAGGGGCCCAATTCATGATCATCATGAA	706
Db	873	CTCCACTTCTATGGGGCTTCTTGATAGGGGCCCAATTCATGATCATCATGAA	932
Qy	707	GCTGTGCGAGGAGAGCTGATGACATACAGAAAGATACCAACAGCATCATTCCTG	766
Db	933	GCTGTGCGAGGAGAGCTGATGACATACAGAAAGATACCAACAGCATCATTCCTG	992
Qy	767	TCTCAAGGCGCTGAGTCTGGGCGCTTTTCTACATGATGGGCTACACATGCTCAG	826
Db	993	TCTCAAGGCGCTGAGTCTGGGCGCTTTTCTACATGATGGGCTACACATGCTCAG	1052
Qy	827	CATCAGAAAGACTATCTCCTCAGCTGAAGACTATGACAAACCCCTTCTGGTCCG	886
Db	1053	CATCAGAAAGACTATCTCCTCAGCTGAAGACTATGACAAACCCCTTCTGGTCCG	1112
Qy	887	CATGTAACGTGATCTGGGGCAAGTTGTGCTGTACAAATATGTCACTGTGGCTGT	946
Db	1113	CATGTAACGTGATCTGGGGCAAGTTGTGCTGTACAAATATGTCACTGTGGCTGT	1172
Qy	947	CACAGAAAGATATGCAATTTGACGGGCGTGGGCTTCAATGGCTTTGAAGAAAGG	1006
Db	1173	CACAGAAAGATATGCAATTTGACGGGCGTGGGCTTCAATGGCTTTGAAGAAAGG	1233
Qy	1007	GGCAAAAGTGGATGCTGTGCAACATGAAAGTGTGCTCTTTGAAACAAACCCCG	1066
Db	1233	GGCAAAAGTGGATGCTGTGCAACATGAAAGTGTGCTCTTTGAAACAAACCCCG	1292
Qy	1067	CACGTGGCAATGTGCTTATCAACATCAACCAACGCTGGGTGGCGCGCTACATTT	1128
Db	1293	CACGTGGCAATGTGCTTATCAACATCAACCAACGCTGGGTGGCGCGCTACATTT	1352
Qy	1127	CAACGACTCAAGTTCCTTGGAAATAAAGACTCTCTCAGGCTCTCTGTGCTATTC	1186

Db 1353 CAAAGACTCAAGTCTCTGGAAAT-AAATACTCTTAAGGCTCTCTGTTGCTATTCCT 1411
Qy 1187 GGCCCTCTGGACGGGCTGCACTCAGGAATPACTGTGTCTGCTTCAGATGGAATTCCTCAT 1246
Db 1412 GGCCCTCTGGACGGGCTGCACTCAGGAATPACTGTGTCTGCTTCAGATGGAATTCCTCAT 1471
Qy 1247 TGTATTTGGAAAAGACAGGCTGCCAGGCTCATTTCAAGAGAGCCCACTGAGCAAGCT 1306
Db 1472 TGTATTTGGAAAAGACAGGCTGCCAGGCTCATTTCAAGAGAGCCCACTGAGCAAGCT 1531
Qy 1307 GGGCGGCAATTAAGTCTCTCCAGCCCTTTACTATTTTGGTGCAACAGACCATTCAGTGGCT 1366
Db 1532 GGGCGGCAATTAAGTCTCTCCAGCCCTTTACTATTTTGGTGCAACAGACCATTCAGTGGCT 1591
Qy 1367 CTTCATGGGTACTTCATGACTGCTCTTCTGCTCAGTGGGAGCAATGGCTTAAGGT 1426
Db 1592 CTTCATGGGTACTTCATGACTGCTCTTCTGCTCAGTGGGAGCAATGGCTTAAGGT 1651
Qy 1427 GATAAATCCATTAATTTCTTGGCCACATCTTCTGAGCCTACTATTCATATTGGC 1486
Db 1652 GATAAATCCATTAATTTCTTGGCCACATCTTCTGAGCCTACTATTCATATTGGC 1711
Qy 1487 TTTATATTCACAAAGCATGTGTCCAAAGAAAGAAAGTTAAGAGATGGAATATTCAT 1546
Db 1712 TTTATATTCACAAAGCATGTGTCCAAAGAAAGAAAGTTAAGAGATGGAATATTCAT 1771
Qy 1547 TTCCTGTGTGGCTGTGTGGGGAATGTGTGAGAAACTGTCTCTTTCACAGCACT 1606
Db 1772 TTCCTGTGTGGCTGTGTGGGGAATGTGTGAGAAACTGTCTCTTTCACAGCACT 1831
Qy 1607 CCTTTGCCCGACAGAGAGATGAAAAGCCAGAGAGGTGGAAGATGCTTCCAGCT 1666
Db 1832 CCTTTGCCCGACAGAGATGAAAAGCCAGAGAGGTGGAAGATGCTTCCAGCT 1891
Qy 1667 GTGCTCTGTGTGCAAGCAAGTCTTATTTGGGCGCAAGAGGAAACTTTTGTGGAG 1726
Db 1892 GTGCTCTGTGTGCAAGCAAGTCTTATTTGGGCGCAAGAGGGAAGC-TTTTGTGGAG 1950
Qy 1727 AGGCGCTTGCTTTGTCAACCACGCTGGAATGCAATGGGCGGATCTCAGCTCACCGCAAC 1786
Db 1951 AGGCGCTTGCTTTGTCAACCACGCTGGAATGCAATGGGCGGATCTCAGCTCACCGCAAC 2010
Qy 1787 CTCACCTCTGTGGGTTCAAGTAATTTTCTGCTCTCAGCCTCCCAAGTGTGGAAATAGA 1846
Db 2011 CTCACCTCTGTGGGTTCAAGTAATTTTCTGCTCTCAGCCTCCCAAGTGTGGAAATAGA 2070
Qy 1847 GGCAGCGCCACATGCCAGCTAATTTTGTATTTTGAAGAAAGGGATTTTCAACACGT 1906
Db 2071 GGCAGCGCCACATGCCAGCTAATTTTGTATTTTGAAGAAAGGGATTTTCAACACGT 2130
Qy 1907 TGGCCAGGCTGTGTGAATCTCTGACCGCAAGTGAATCCACCGGCTCCGCCCTCCCAAG 1966
Db 2131 TGGCCAGGCTGTGTGAATCTCTGACCGCAAGTGAATCCACCGGCTCCGCCCTCCCAAG 2190
Qy 1967 TGTGGGATTAACAGGGGTGAGCCAGCGTGCAGGCGCCCAAGGGGAAACTTTGTGGAG 2026
Db 2191 TGTGGGATTAACAGGGGTGAGCCAGCGTGCAGGCGCCCAAGGGGAAACTTTGTGGAG 2250
Qy 2027 AGCAGAGGGGCTCAATCTCCCTGTGATTTCCCAATGCAATGCTTATCTTCCCA 2086
Db 2251 AGCAGAGGGGCTCAATCTCCCTGTGATTTCCCAATGCAATGCTTATCTTCCCA 2310
Qy 2087 TCTAGCAGGAATCTATTTGTTTCTTCTGCAATTTAATTAATGTAATGTGGCG 2146
Db 2311 TCTAGCAGGAATCTATTTGTTTCTTCTGCAATTTAATTAATGTAATGTGGCG 2370
Qy 2147 CTACACCAACCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCTGTGCTGCACCT 2206
Db 2371 CTACACCAACCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCTGTGCTGCACCT 2430
Qy 2207 TTTCTACCTTGAAGCTGTATTAATAATCACTTCTGTTTGTACGTTTT 2258
Db 2431 TTTCTACCTTGAAGCTGTATTAATAATCACTTCTGTTTGTACGTTTT 2482

RESULT 5
AX31926
LOCUS AX31926 1842 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2435 from Patent WO0194629.
ACCESSION AX31926
VERSION AX31926.1 GI:18122560
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Hom sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ehner, R., Endress, G.,
Horrigan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 2435 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
location/Qualifiers
1..1842
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 413 a 525 c 413 g 491 t
ORIGIN
Query Match 81.0%; Score 1842; DB 6; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 279 TACCTCATCACCCTCTTCATACCTTTAAGGCTCTCAATGCTATTTTAAGCTTTGA 338
Db 1 TACCTCATCACCCTCTTCATACCTTTAAGGCTCTCAATGCTATTTTAAGCTTTGA 60
Qy 339 AACCACTTACACCTCCCTGTGTGATTTGTCTCACTTCTCATCTTCCAGTAAAG 398
Db 61 AACCACTTACACCTCCCTGTGTGATTTGTCTCACTTCTCATCTTCCAGTAAAG 120
Qy 399 GGCCGACCATCACTGCGCTCTCACTACCTTTTGGCTTCCAGATGGCTTCTGCT 458
Db 121 GGCCGACCATCACTGCGCTCTCACTACCTTTTGGCTTCCAGATGGCTTCTGCT 180
Qy 459 GGATCTATTAACGTGCGCAACCGCAACTAAGATCAAGTGAATGCAATGCAATGTGT 518
Db 181 GGATCTATTAACGTGCGCAACCGCAACTAAGATCAAGTGAATGCAATGCAATGTGT 240
Qy 519 CTGACTTTGAACCTGATTTGTTGGCTGTGACTCTTTGAAGAGGAAATCAGAAAT 578
Db 241 CTGACTTTGAACCTGATTTGTTGGCTGTGACTCTTTGAAGAGGAAATCAGAAAT 300
Qy 579 TCTTGTCTCTGAGCAACAGAAATATGCAATAGTGTGTTCTTCTCTGCTGAAGTT 638
Db 301 TCTTGTCTCTGAGCAACAGAAATATGCAATAGTGTGTTCTTCTCTGCTGAAGTT 360
Qy 638 GCTGTTTCTCTACTTCTTAATGGGCTCTTGTGAGAGGCCCAAGTTCATGAAATC 698
Db 361 GCTGTTTCTCTACTTCTTAATGGGCTCTTGTGAGAGGCCCAAGTTCATGAAATC 420
Qy 699 TACATGAAGCTGTGAGAGGAGAGTGAATGATACACAGAAAGATACCAACAGATC 758
Db 421 TACATGAAGCTGTGAGAGGAGAGTGAATGATACACAGAAAGATACCAACAGATC 480
Qy 759 ATTCTGCTCTGAAGCGCTGAGTGTGGGCTTTTCTACTGATGAGGCTACACATGCTC 818
Db 481 ATTCTGCTCTGAAGCGCTGAGTGTGGGCTTTTCTACTGATGAGGCTACACATGCTC 540
Qy 819 AACCCCAATACAGAAAGATATCTCTCACTGAAGACTATGACAAACCAACCCCTTCTG 878
Db 541 AACCCCAATACAGAAAGATATCTCTCACTGAAGACTATGACAAACCAACCCCTTCTG 600
Qy 879 TTCGCTGATGATGATGCTGATCTGGGCAAGTTTGTCTGTATACAAATATGACCTGT 938

Db 601 TTCCGCTGACATGATGATCTGGGCAAGTTGTGCTGTAACAATATGTCACCTGT 660
Qy 939 TGGCTGGTCAAGAGAGATATGATTTTGAACGGGCTGGGCTTCAATGGCTTTGAAGA 998
Db 661 TGGCTGGTCAAGAGAGATATGATTTTGAACGGGCTGGGCTTCAATGGCTTTGAAGA 720
Qy 999 AAGGGCAAGCAAGTGGGATGCTGTGCCAATGAGGTGGCTCTTTGAACAAC 1058
Db 721 AAGGGCAAGCAAGTGGGATGCTGTGCCAATGAGGTGGCTCTTTGAACAAC 780
Qy 1059 CCCCCTTCACTGGACACATGCTCTCATTTCAATCAACCAAGGCTGGGCCCCG 1118
Db 781 CCCCCTTCACTGGACACATGCTCTCATTTCAATCAACCAAGGCTGGGCCCCG 840
Qy 1119 TACATCTTCAACGACATCAAGTCTTGAATAATTAAGACTCTCTCAGGGCTCTCGTT 1178
Db 841 TACATCTTCAACGACATCAAGTCTTGAATAATTAAGACTCTCTCAGGGCTCTCGTT 900
Qy 1179 CTATTCCTGGCCCTCTGGCAAGGCTGCACTCAGGATACCTGTCTGCTTCCAGATGAA 1238
Db 901 CTATTCCTGGCCCTCTGGCAAGGCTGCACTCAGGATACCTGTCTGCTTCCAGATGAA 960
Qy 1239 TTCCCTATTTGATTTGGAAGAGAGAGGCTGCAAGGCTCAATTCAGAGAGCCCACTG 1298
Db 961 TTCCCTATTTGATTTGGAAGAGAGAGGCTGCAAGGCTCAATTCAGAGAGCCCACTG 1020
Qy 1299 AGCAAGCTGGCCGCAATTAATGCTCTCAGGCTTCTACTATTTGTTGCAACAGACATC 1358
Db 1021 AGCAAGCTGGCCGCAATTAATGCTCTCAGGCTTCTACTATTTGTTGCAACAGACATC 1080
Qy 1359 CACTGGCTCTTCAATGGGTTACTCATGATGCTCTTCTGCTCTTCACTGGGCAATG 1418
Db 1081 CACTGGCTCTTCAATGGGTTACTCATGATGCTCTTCTGCTCTTCACTGGGCAATG 1140
Qy 1419 CTTAAGGTGATTAATCATATATTTCTGGGCAACATCTTCTGAGCTTACTATTC 1478
Db 1141 CTTAAGGTGATTAATCATATATTTCTGGGCAACATCTTCTGAGCTTACTATTC 1200
Qy 1479 ATATTCCTTATATTCACAAAGCAATGTCGCAAGAAAGAAAGTTAAAGAGATGAA 1538
Db 1201 ATATTCCTTATATTCACAAAGCAATGTCGCAAGAAAGAAAGTTAAAGAGATGAA 1260
Qy 1539 TAATTCATTTCTGCTGGGCTGTGCGGAGCTGTGCAAGAACTACTGCTCTCTTTTC 1598
Db 1261 TAATTCATTTCTGCTGGGCTGTGCGGAGCTGTGCAAGAACTACTGCTCTCTTTTC 1320
Qy 1599 ACACACCTCTTGGCCCAAGAGAGAGAAAGGAGGAGGAGATGATGATG 1658
Db 1321 ACACACCTCTTGGCCCAAGAGAGAGAAAGGAGGAGGAGATGATGATG 1380
Qy 1659 TTCCAGCTGTGCTCTGTGTCAGCAAGTCTTCAATTTGGGCAAGAGGGAATCTTTT 1718
Db 1381 TTCCAGCTGTGCTCTGTGTCAGCAAGTCTTCAATTTGGGCAAGAGGGAATCTTTT 1440
Qy 1719 TTTGGAAGAGGCTGTGCTTTGTACCCACGCTGGAATGCAAGTGGGGATCTCAGCTC 1778
Db 1441 TTTGGAAGAGGCTGTGCTTTGTACCCACGCTGGAATGCAAGTGGGGATCTCAGCTC 1500
Qy 1779 ACCGCACTCTCAGCTCTGCTGGGTTCAAGTATTTTCCGCTCAGGCTCCCAAGTATG 1838
Db 1501 ACCGCACTCTCAGCTCTGCTGGGTTCAAGTATTTTCCGCTCAGGCTCCCAAGTATG 1560
Qy 1839 GGAATACAGGACGCAACATGCTCAGCTAATTTTGTATTTTCAAGTAAACGGGATTT 1898
Db 1561 GGAATACAGGACGCAACATGCTCAGCTAATTTTGTATTTTCAAGTAAACGGGATTT 1620
Qy 1899 CACCACTTTGGCCAGGCTGTGCTGAACTCTGACCCGAATGATCCACCCGCTCGCC 1958
Db 1621 CACCACTTTGGCCAGGCTGTGCTGAACTCTGACCCGAATGATCCACCCGCTCGCC 1680
Qy 1959 TCCCAAGTGTGGGATTAAGAGGCTGAGCAAGCTGCTCGGCAAGGGGGAATCTTT 2018

Db 1681 TCCCAAGTGTGGGATTAAGAGGCTGAGCAAGCTGCTCGGCAAGGGGAACTCTT 1740
Qy 2019 GTGGAGAGACAGAGGGGCTCAATCTCCCTCTGATTTCCCAAGACATGCTTATC 2078
Db 1741 GTGGAGAGACAGAGGGGCTCAATCTCCCTCTGATTTCCCAAGACATGCTTATC 1800
Qy 2079 TCTCCCATCTAGCCAGGAATCTATGTTTCTTTCTTCCGCC 2120
Db 1801 TCTCCCATCTAGCCAGGAATCTATGTTTCTTTCTTCCGCC 1842

RESULT 6
AX410769 1842 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 3416 from Patent WO0229103.
DEFINITION AX410769
ACCESSION AX410769
VERSION AX410769.1 GI:2143474
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 3416 11-APR-2002;
GENE LOGIC INC (US)
Location/Qualifiers
FEATURES
source
1..1842
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. U72515"
BASE COUNT 413 a 525 c 413 g 491 t
ORIGIN

Query Match 81.0%; Score 1842; DB 6; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 TACCTATCCACCTCTTCCATACCTTTTACAGGCTCTCAATGCTTATTTTAACTTTGA 338
Db 1 TACCTATCCACCTCTTCCATACCTTTTACAGGCTCTCAATGCTTATTTTAACTTTGA 60
Qy 339 AACCACTTACCACTCTGCTGTGTATTTGCTTCACTTCTCATCTTGAATGATG 398
Db 61 AACCACTTACCACTCTGCTGTGTATTTGCTTCACTTCTCATCTTGAATGATG 120
Qy 399 GGGCGACATCACTGCGCTCTCACTCTTTTGTCCAGATGGCTTACCTTGGCT 458
Db 121 GGGCGACATCACTGCGCTCTCACTCTTTTGTCCAGATGGCTTACCTTGGCT 180
Qy 459 GGATATTTTACACTGCGCAACGCACTAGATATCAAGTGAAGATGCAATTTGTT 518
Db 181 GGATATTTTACACTGCGCAACGCACTAGATATCAAGTGAAGATGCAATTTGTT 240
Qy 519 CTGACTTTGAAGCTGATTTGCTGTGTTGCTTACTTTTGAAGAGGAAAGATCAGAT 578
Db 241 CTGACTTTGAAGCTGATTTGCTGTGTTGCTTACTTTTGAAGAGGAAAGATCAGAT 300
Qy 579 TCCCTGTCTCTAGCAACGAATATGCAATGCGATGCGTCTTCCCTGCTGGAAGTT 638
Db 301 TCCCTGTCTCTAGCAACGAATATGCAATGCGATGCGTCTTCCCTGCTGGAAGTT 360
Qy 639 GCTGTTTCTCTACTTCTATTTGAGGCTCTTGTGTAAGGCTCCAGATTTCTCAATGAATCAC 698
Db 361 GCTGTTTCTCTACTTCTATTTGAGGCTCTTGTGTAAGGCTCCAGATTTCTCAATGAATCAC 420
Qy 699 TACATGAAGTGTGCAAGGAGAGCTGATTTGATATACATACAGAAAGATACCAACAGATC 758
Db 421 TACATGAAGTGTGCAAGGAGAGCTGATTTGATATACATACAGAAAGATACCAACAGATC 480
Qy 759 ATTCTGTCTCAAGGCTCTGAGTCTGGGCTTTTCTACTAGTGGGCTACACATGCTCT 818

```

Db      481  |||||ATTCTGCTCTCAAGGCGCTGAGCTGGGCGCTTTTCTACTAGGGCTACACATGCTC 540
Qy      819  |||||AGCCCCACATCAGAGAGCTATCTCTCTACGGAAGACTATAGAACACACCCCTTCTG 878
Db      541  |||||AGCCCCACATCAGAGAGCTATCTCTCTACGGAAGACTATAGAACACACCCCTTCTG 600
Qy      879  |||||TTCGGTGCATGTACATGTGATCTGGGGCAAGTTTGTGCTGTACAAATATGTCACTGT 938
Db      601  |||||TTCGGTGCATGTACATGTGATCTGGGGCAAGTTTGTGCTGTACAAATATGTCACTGT 660
Qy      939  |||||TGGCTGTCAAGAGAGATATGCAATTTTGAACGGGCTTGAAGTTTGAAGAA 998
Db      661  |||||TGGCTGTCAAGAGAGATATGCAATTTTGAACGGGCTTGAAGTTTGAAGAA 720
Qy      999  |||||AAGGGAAGGCAAGTGGGATGCTGTGCGCAATAGAGGTGGCTTTTGAACAAAC 1058
Db      721  |||||AAGGGAAGGCAAGTGGGATGCTGTGCGCAATAGAGGTGGCTTTTGAACAAAC 780
Qy      1059 |||||CCCCGTTCACTGGCAACATTCCTCATTCATCAATCAACACACGCTTGGGCGCG 1118
Db      781  |||||CCCCGTTCACTGGCAACATTCCTCATTCATCAATCAACACACGCTTGGGCGCG 840
Qy      1119 |||||TACATCTTCAAGACTCAAGTTCTTGGAAATTAAGAACTCTCTAGGGTCTCTGCTG 1178
Db      841  |||||TACATCTTCAAGACTCAAGTTCTTGGAAATTAAGAACTCTCTAGGGTCTCTGCTG 900
Qy      1179 |||||CTATTCTGGCGCTCTGGGACGGGCGCTGCACTGAGATACCTGCTCTTCCAGATGGA 1238
Db      901  |||||CTATTCTGGCGCTCTGGGACGGGCGCTGCACTGAGATACCTGCTCTTCCAGATGGA 960
Qy      1239 |||||TTCCTCATTTGTTATTTGGAAAGACAGGCTGCCAGGCTCATTAAGAGAGCCCCACTG 1298
Db      961  |||||TTCCTCATTTGTTATTTGGAAAGACAGGCTGCCAGGCTCATTAAGAGAGCCCCACTG 1020
Qy      1299 |||||AGCAAGCTGGCGCCATTAATCTGCTCCAGCGCTTCTACTATTTGGTGCAACAGACATC 1358
Db      1021 |||||AGCAAGCTGGCGCCATTAATCTGCTCCAGCGCTTCTACTATTTGGTGCAACAGACATC 1080
Qy      1359 |||||CACTGGCTCTTCAATGGGTTAATCCATGACGCTCTTGGCTCTTCAAGTGGGCAATAG 1418
Db      1081 |||||CACTGGCTCTTCAATGGGTTAATCCATGACGCTCTTGGCTCTTCAAGTGGGCAATAG 1140
Qy      1419 |||||CTTAAGGTATTAATCAATCTATTTCTTGGGCAATCTTCTCTGAGGCTACTATTC 1478
Db      1141 |||||CTTAAGGTATTAATCAATCTATTTCTTGGGCAATCTTCTCTGAGGCTACTATTC 1200
Qy      1479 |||||ATATTGCTTTATTAATCAAGCAATGTGCCAAGGAAAGAGAAATTAAGAGATGGA 1538
Db      1201 |||||ATATTGCTTTATTAATCAAGCAATGTGCCAAGGAAAGAGAAATTAAGAGATGGA 1260
Qy      1539 |||||TAAATCAATTTCCCTGGTGGCTGTGGGAGCTGTGTGCAAGAACTACTCTCTCTTTTC 1598
Db      1261 |||||TAAATCAATTTCCCTGGTGGCTGTGGGAGCTGTGTGCAAGAACTACTCTCTCTTTTC 1320
Qy      1599 |||||ACAGCACTCTTTTCCCGAGAGAGAGAAATGAAAGCCAGGAGGTGGAAGATCGATGC 1658
Db      1321 |||||ACAGCACTCTTTTCCCGAGAGAGAGATGAAAGCCAGGAGGTGGAAGATCGATGC 1380
Qy      1659 |||||TTCAGCTGTGCTCTGTGCGCAGCAAGCTTCAATTTGGGGCAAGGGGAAACTTTT 1718
Db      1381 |||||TTCAGCTGTGCTCTGTGCGCAGCAAGCTTCAATTTGGGGCAAGGGGAAACTTTT 1440
Qy      1719 |||||TTTGGAGAGGGCTCTTGTCTTGTCAACACGCTGGAATGCAATGCGGGATCTCAGCTC 1778
Db      1441 |||||TTTGGAGAGGGCTCTTGTCTTGTCAACACGCTGGAATGCAATGCGGGATCTCAGCTC 1500
Qy      1779 |||||ACCGCAACTCCACTCTGCTGGGTTCAAGTATTTTCTGCTCAGGCTCCCAAGTAGCTG 1838
Db      1501 |||||ACCGCAACTCCACTCTGCTGGGTTCAAGTATTTTCTGCTCAGGCTCCCAAGTAGCTG 1560
Qy      1839 |||||GGATTCAGGCAAGCCACCAATGCCAGCTAATTTTGTATTTTCAATAGAAACGGGATTT 1898

```

```

Db      1561 GGAATACAGGACAGCCACCATCCAGCTAATTTTGTATTTTCAATAGAAACGGGATTT 1620
Qy      1899  |||||CACCAAGTTGGCCAGGCTGTGTCTCGAATCTCTGACCGCAAGTATCCACCCGCTCCGC 1958
Db      1621  |||||CACCAAGTTGGCCAGGCTGTGTCTCGAATCTCTGACCGCAAGTATCCACCCGCTCCGC 1680
Qy      1959  |||||TCCCAAGTGTGGGATTAACAGGCGTGAAGCCACCGGCGCCCAAGGGGAACTTT 2018
Db      1681  |||||TCCCAAGTGTGGGATTAACAGGCGTGAAGCCACCGGCGCCCAAGGGGAACTTT 1740
Qy      2019  |||||GTGGAGAGAGAGAGGGGCTCAATCTCCCTCTGATTTCCCATGACATTCCTTATC 2078
Db      1741  |||||GTGGAGAGAGAGAGGGGCTCAATCTCCCTCTGATTTCCCATGACATTCCTTATC 1800
Qy      2079  |||||TCTCCCATCTAGCCAGGAATCTATGTGTTTTTTCTTCTGCC 2120
Db      1801  |||||TCTCCCATCTAGCCAGGAATCTATGTGTTTTTTCTTCTGCC 1842

RESULT 7
HSU72515
LOCUS      HSU72515      1842 bp      mRNA      linear      PRI 24-JUL-1997
DEFINITION Human C3f mRNA, complete cds.
ACCESSION U72515
VERSION   U72515.1 GI:1673519
KEYWORDS
SOURCE
ORGANISM  Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
  1 (bases 1 to 1842)
    Ansari-Lari, M.A., Shen, Y., Muzny, D.M., Lee, W. and Gibbs, R.A.
    Large-scale sequencing in human chromosome 12p13: experimental and
    computational gene structure determination
    Genome Res. 7 (3), 268-280 (1997)
JOURNAL   97228904
MEDLINE   9074930
REFERENCE
  2 (bases 1 to 1842)
    Ansari-Lari, M.A., Shen, Y., Muzny, D.M., Lee, W. and Gibbs, R.A.
    Direct Submission
    Submitted (24-SEP-1996) Molecular and Human Genetics, Baylor
    College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
FEATURES
    source
        1..1842
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="12"
            /map="12p13"
            1..1842
            /gene="C3f"
            1..1842
            /gene="C3f"
            118..1263
            /gene="C3f"
            /note="similar to ESTs with GenBank Accession Numbers
            H45806 and H45837; similar to S. cerevisiae ORF YOR175c
            encoded by GenBank Accession Number Z75083; see
            corresponding genomic sequence in GenBank Accession Number
            U72506"
            /codon_start=1
            /product="C3f"
            /protein_id="AAC51640.1"
            /db_xref="GI:1673520"
            /translation="MGRITVALLTFCFQMAVLLAGYYTATGNDIKMTMPCVLT
            KLGIADVDFGKDONLSBOQKYLARGVSLLEVAFSYFGAFLVGPQPSNMHY
            MKLVGGLIDIPGIPNSIIPALKRLSGLFVLVGTLSLPHITEDYLTEDYDHPF
            MPRCYMLIMGKFPVLKYVTCMLVTEGVCILTGAGNPEEKAKMPCANAKMYLF
            ETNPRFTIASFNINTNANVARYIFKRLKPLKNEKLSOGSLFLALRHGHSYLV
            CFQEFLLVVERQARLIQSPSTLKLAAITVLDPFYLLVOOTIHMVPMGYSMTAFC
            LFTMDKMLKVKYSIYFGLHIFLSLFLPIYIHKMVPKREKLRKME"
BASE COUNT      413 a      525 c      413 g      491 t
ORIGIN

```


Query Match 81.0%; Score 1842; DB 9; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

279 TACCTATCCACCTCTTCCATCTTATGAGGCTCTCATCTGTTATTTTAACTTTGA 338
1 TACTCATCCACCTCTTCCATCTTATGAGGCTCTCATCTGTTATTTTAACTTTGA 60

339 AACGAGCTTACACCTCTGCTGATGATGCTGATGCTGATGCTGATGCTGATG 338
61 AACGAGCTTACACCTCTGCTGATGATGCTGATGCTGATGCTGATGCTGATG 120

399 GGGCGACCATCACTGCGCTCTCACTACCTTTTGTCTTCAGATGCTTACCTT 458
121 GGGCGACCATCACTGCGCTCTCACTACCTTTTGTCTTCAGATGCTTACCTT 180

459 GGATACATTTACATGCGACCGGCACTAGATATCAAGTGGAGCAATGGCACTT 518
181 GGATACATTTACATGCGACCGGCACTAGATATCAAGTGGAGCAATGGCACTT 240

519 CTGACTTTGAAGCTGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 578
241 CTGACTTTGAAGCTGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300

579 TCTCTGCTCTGAGCAACAGAAATATGCCATACGTGGTGTCTCTCTCTGGA 638
301 TCTCTGCTCTGAGCAACAGAAATATGCCATACGTGGTGTCTCTCTCTGGA 360

639 GCTGGTTTCT 698
361 GCTGGTTTCT 420

699 TACATGAAGCTGCTGAGGAGAGCTGATGATGATGATGATGATGATGATGAT 758
421 TACATGAAGCTGCTGAGGAGAGCTGATGATGATGATGATGATGATGATGAT 480

759 ATTCTGCTCTCAAGCGCTGATGCTGAGGCTTTTCTTACTGATGCTGATG 818
481 ATTCTGCTCTCAAGCGCTGATGCTGAGGCTTTTCTTACTGATGCTGATG 540

819 AGCCCAACATCAAGAACTATCTCTCTGATGATGATGATGATGATGATGAT 878
541 AGCCCAACATCAAGAACTATCTCTCTGATGATGATGATGATGATGATGAT 600

879 TTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 938
601 TTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

939 TGGCTGCTCAAGAAAGATGATGATGATGATGATGATGATGATGATGATG 998
661 TGGCTGCTCAAGAAAGATGATGATGATGATGATGATGATGATGATGATG 720

999 AAGGCAAGGCAAGTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
721 AAGGCAAGGCAAGTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

1059 CCCGCTTCACTGCGACCATGCTCTCATTTCAATCAACACCAACGCTGCT 1118
781 CCCGCTTCACTGCGACCATGCTCTCATTTCAATCAACACCAACGCTGCT 840

1119 TACATCTTCAAGAGCTCAAGTCTCTTGGAAATAAGAACTCTCAAGGCTCT 1178
841 TACATCTTCAAGAGCTCAAGTCTCTTGGAAATAAGAACTCTCAAGGCTCT 900

1179 CTATTCCTGCGCTCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1238
901 CTATTCCTGCGCTCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

1239 TTCTCATTTGTTATTTGGAAGAAGAGGCTCCAGGCTCATTTCAAGAGGCC 1298
961 TTCTCATTTGTTATTTGGAAGAAGAGGCTCCAGGCTCATTTCAAGAGGCC 1020

QY 1299 AGCAAGCTGGCGCCCATTAATGCTCTTCCAGCCCTTACTATTTGGTGAACAGACATC 1358
DB 1021 AGCAAGCTGGCGCCCATTAATGCTCTTCCAGCCCTTACTATTTGGTGAACAGACATC 1080

QY 1359 CACTGGCTCTTCAATGAGTACTCATGATGATGATGATGATGATGATGATGAT 1418
DB 1081 CACTGGCTCTTCAATGAGTACTCATGATGATGATGATGATGATGATGATGAT 1140

QY 1419 CTTAAGGTATTAATCAATCTATTTCTTGGCACATCTTCTCTGAGCTTACTATTC 1478
DB 1141 CTTAAGGTATTAATCAATCTATTTCTTGGCACATCTTCTCTGAGCTTACTATTC 1200

QY 1479 ATATGCTTATTAATTAAGAAAGATGATGATGATGATGATGATGATGATGAT 1538
DB 1201 ATATGCTTATTAATTAAGAAAGATGATGATGATGATGATGATGATGATGAT 1260

QY 1539 TAATCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1598
DB 1261 TAATCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

QY 1599 ACAGCACTCTTGGCCAGAGAGATGATGATGATGATGATGATGATGATGAT 1658
DB 1321 ACAGCACTCTTGGCCAGAGAGATGATGATGATGATGATGATGATGATGAT 1380

QY 1659 TTCCAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1718
DB 1381 TTCCAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440

QY 1719 TTGGAGAAAGCGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1778
DB 1441 TTGGAGAAAGCGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1500

QY 1779 ACCGCAACTCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1838
DB 1501 ACCGCAACTCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560

QY 1839 GGAATACAGGACAGGACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1898
DB 1561 GGAATACAGGACAGGACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620

QY 1899 CACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1958
DB 1621 CACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680

QY 1959 TCCCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2018
DB 1681 TCCCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

QY 2019 GTGGAGAGAGAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2078
DB 1741 GTGGAGAGAGAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800

QY 2079 TCTCCCATCTAGCAGGAATCTATTTGTTTCTTCTGCTC 2120
DB 1801 TCTCCCATCTAGCAGGAATCTATTTGTTTCTTCTGCTC 1842

RESULT 8
AX073560 1461 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 12 from Patent WO0104297.
ACCESSION AX073560
VERSION AX073560.1 GI:12709973
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Kato, S. and Kimura, T.
Human proteins having hydrophobic domains and dnas encoding these
proteins
JOURNAL Patent: WO 0104297-A 12 18-JAN-2001;

SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)
 Location/Qualifiers
 1.1461
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 326 a 401 c 336 g 398 t
 ORIGIN
 Query Match 64.3%; Score 1461; DB 6; Length 1461;
 Best Local Similarity 100.0%; Pred No. 0;
 Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 78 ATGGCGTCTCAGCGAGGGGAGCGAGGGAGCTGTGGTGGCGGTCTGCGAG 137
 1 ATGGCGTCTCAGCGAGGGGAGCGAGGGAGCTGTGGTGGCGGTCTGCGAG 60
 138 TGGGTTTCAGAGAGCTGAGCTTAAAGAGTTGGCGAGTCCCTGGGCGGTGAGAG 197
 61 TGGGTTTCAGAGAGCTGAGCTTAAAGAGTTGGCGAGTCCCTGGGCGGTGAGAG 120
 198 GGGCTGGGCTGATCATCTCCATCTCCGGGTTACCCCTTGTGTTTATGGGAT 257
 121 GGGCTGGGCTGATCATCTCCATCTCCGGGTTACCCCTTGTGTTTATGGGAT 180
 258 TACCTTTTACAGAGAGAGCTACCTCATCACTCTTCATACCTTTACAGGCTCTCA 317
 181 TACCTTTTACAGAGAGAGCTACCTCATCACTCTTCATACCTTTACAGGCTCTCA 240
 318 ATTGCTATTATTTAATCTTTGGAAACAGCTCTACCACTCCCTGTGTGTAATGTCTGAG 377
 241 ATTGCTATTATTTAATCTTTGGAAACAGCTCTACCACTCCCTGTGTGTAATGTCTGAG 300
 378 TTCTCATCTCTGAGATAATGGGCGGACATCACTGCGGCTCTCACTACCTTTGCTTC 437
 301 TTCTCATCTCTGAGATAATGGGCGGACATCACTGCGGCTCTCACTACCTTTGCTTC 360
 438 CAGATGGCTCACTCTCTGCTGATATTAATCACTGCGGAGCGGAGATATCAAG 497
 361 CAGATGGCTCACTCTCTGCTGATATTAATCACTGCGGAGCGGAGATATCAAG 420
 498 TGGACATGCCACATTTGTTCTGACTTTGAAGCTGATTTGGCTGTGTAATCACTTT 557
 421 TGGACATGCCACATTTGTTCTGACTTTGAAGCTGATTTGGCTGTGTAATCACTTT 480
 558 GACGGAGGGAAGATAGAAATCTTGTCTCTGAGCAACAGAAATATGCCATAGCTGT 617
 481 GACGGAGGGAAGATAGAAATCTTGTCTCTGAGCAACAGAAATATGCCATAGCTGT 540
 618 GTTCCCTCCCTGCTGGAAGTTGCTGTTCTCTTAATTTATGGGCTTTCTTGTGAGG 677
 541 GTTCCCTCCCTGCTGGAAGTTGCTGTTCTCTTAATTTATGGGCTTTCTTGTGAGG 600
 678 CCCCAATCTCAATGATCATCAATGAGCTGTGGAGGAGAGCTGATGATGACATGCA 737
 601 CCCCAATCTCAATGATCATCAATGAGCTGTGGAGGAGAGCTGATGATGACATGCA 660
 738 GGAAGATATCCAAAGCATCATTTCTGCTTCAAGCGCTGAGTGTGGGCTTTTCTAC 797
 661 GGAAGATATCCAAAGCATCATTTCTGCTTCAAGCGCTGAGTGTGGGCTTTTCTAC 720
 798 CTAGTGGCTTACAGACTGCTGAGCCGCCACATCAAGAGACTATCTCTCACTGAAGC 857
 721 CTAGTGGCTTACAGACTGCTGAGCCGCCACATCAAGAGACTATCTCTCACTGAAGC 780
 858 TATGACACACCCCTTCTGTGTTCCGCTGATGATGATGATGATGATGATGATGATG 917
 781 TATGACACACCCCTTCTGTGTTCCGCTGATGATGATGATGATGATGATGATGATG 840
 918 CTGTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 977
 841 CTGTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900

QY 978 GGCCTCAATGCTTTGAAGAAAAGGCAAGGCAAGATGAGATGCTGTGCCAATGAG 1037
 DB 901 GGCCTCAATGCTTTGAAGAAAAGGCAAGGCAAGATGAGATGCTGTGCCAATGAG 960
 QY 1038 GTGTGGCTTTTGAAGAAAAGGCAAGGCAAGATGAGATGCTGTGCCAATGAG 1097
 DB 961 GTGTGGCTTTTGAAGAAAAGGCAAGGCAAGATGAGATGCTGTGCCAATGAG 1020
 QY 1098 ACCAAGGCTGGGAGGCGGCTGACATCTTCAAGAGATGAGATGCTGTGCCAATGAG 1157
 DB 1021 ACCAAGGCTGGGAGGCGGCTGACATCTTCAAGAGATGAGATGCTGTGCCAATGAG 1080
 QY 1158 CTCTCTCAAGGCTCTCTGTGTGATTTCTGAGGCTCTGTGAGGCTGTGAGGCTGTGAG 1217
 DB 1081 CTCTCTCAAGGCTCTCTGTGTGATTTCTGAGGCTCTGTGAGGCTGTGAGGCTGTGAG 1140
 QY 1218 CTGTGTGCTTCAAGTGAATTTCTCATTTGTTATTTGAGAAAGACAGGCTGCCAGGCTC 1277
 DB 1141 CTGTGTGCTTCAAGTGAATTTCTCATTTGTTATTTGAGAAAGACAGGCTGCCAGGCTC 1200
 QY 1278 ATTCAAGAGAGGCGGCAAGGCTGAGAGAGGCGGCTGAGGCTGAGGCTGAGGCTGAG 1337
 DB 1201 ATTCAAGAGAGGCGGCAAGGCTGAGAGAGGCGGCTGAGGCTGAGGCTGAGGCTGAG 1260
 QY 1338 TATTTGTCGCAACAGACCATGCTGCTCTTCAATGAGGTTACTGATGATGCTGCTTCTG 1397
 DB 1261 TATTTGTCGCAACAGACCATGCTGCTCTTCAATGAGGTTACTGATGATGCTGCTTCTG 1320
 QY 1398 CTCTTCAAGGCGGCAAGGCTGAGAGAGGCTGAGAGAGGCTGAGAGAGGCTGAGAGAG 1457
 DB 1321 CTCTTCAAGGCGGCAAGGCTGAGAGAGGCTGAGAGAGGCTGAGAGAGGCTGAGAGAG 1380
 QY 1458 TTCTTCTGAGGCTCTATCATTTGCTTATTTCAAGGCTGAGAGAGGCTGAGAGAG 1517
 DB 1381 TTCTTCTGAGGCTCTATCATTTGCTTATTTCAAGGCTGAGAGAGGCTGAGAGAG 1440
 QY 1518 GAGAGTTAAAGAGATGAGAA 1538
 DB 1441 GAGAGTTAAAGAGATGAGAA 1461

RESULT 9
 AK058063 1573 bp mRNA linear PRI 31-OCT-2001
 LOCUS
 DEFINITION
 AK058063 Homo sapiens cDNA FLJ25334 fis, clone TS100701.
 VERSION
 AK058063.1 GI:16554077
 KEYWORDS
 oligo cloning; fis (full insert sequence).
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1
 Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
 Ono, Y., Horiuchi, T., Hiraoka, S., Murakawa, K., Takiguchi, S.,
 Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
 Yamashita, H., Chiba, Y., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S.,
 Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
 Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isegai, T. and
 Sugano, S.
 NEDO human cDNA sequencing project
 TITLE
 NEDO human cDNA sequencing project
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 1573)
 AUTHORS
 Sugano, S. and Suzuki, Y.
 TITLE
 Direct Submision
 JOURNAL
 Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology (RAB); cDNA library

construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; Clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES
source
1. .1573
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TST0701"
/issue_type="testis"
/clone_lib="TST"
/note="Cloning vector: pME18SFL3"

BASE COUNT 358 a 436 c 378 g 401 t

ORIGIN

Query Match 54.5%; Score 1238; DB 9; Length 1573;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1388; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

864 AACACCCCTTCTGCTCCGCTGATGATGCTGATCTGGGGCAAGTTGCTGTAC 923
183 AACACCCCTTCTGCTCCGCTGATGATGCTGATCTGGGGCAAGTTGCTGTAC 242
924 AAATATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
243 AAATATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302
984 AATGCTTTAAGAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1043
303 AATGCTTTAAGAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 362
1044 CTCTTTGAAACAAACCCCGCTTCACTGCGACCATGCTCTTCAATCAACCAAC 1103
363 CTCTTTGAAACAAACCCCGCTTCACTGCGACCATGCTCTTCAATCAACCAAC 422
1104 GCGTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163
423 GCGTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
1164 CAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1223
483 CAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
1224 TGCTTCAAGATGATTCCTCATTTGTTATGTTGAAAGACAGCTGCCAGCTCAT 1283
543 TGCTTCAAGATGATTCCTCATTTGTTATGTTGAAAGACAGCTGCCAGCTCAT 602
1284 GAGAGCCCAACCCGAGAGAGCTGGCGCATTACTGCTCCGAGCCCTTCAATTTG 1343
603 GAGAGCCCAACCCGAGAGAGCTGGCGCATTACTGCTCCGAGCCCTTCAATTTG 662
1344 GTGCAAGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403
663 GTGCAAGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
1404 ACGTGGACAAATGGCTTAAGGTGTAATATCAATTAATTTCTTGGCCATCTTCT 1463
723 ACGTGGACAAATGGCTTAAGGTGTAATATCAATTAATTTCTTGGCCATCTTCT 782
1464 CTGAGGCTATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1523
783 CTGAGGCTATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
1524 TTAAGAAGATGATATCAATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1583
843 TTAAGAAGATGATATCAATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
1584 CTGCTCTCTCTTTTCAAGCACTCTCTTGGCCCAAGAGACAGAAATGCAAGGAG 1643
903 CTGCTCTCTCTTTTCAAGCACTCTCTTGGCCCAAGAGACAGAAATGCAAGGAG 962
1644 GTGAAGATGATGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1703

1704 AAGGGGAACTTTTGTGAGAGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1763
1023 AAGGGGAACTTTTGTGAGAGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1082
1764 GCGGATCTCAGCTACCGCCACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1823
1083 GCGGATCTCAGCTACCGCCACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142
1824 CTCCCAAGATGCTGGAATATACAGGACGACCATGCTGCTGCTGCTGCTGCTG 1883
1143 CTCCCAAGATGCTGGAATATACAGGACGACCATGCTGCTGCTGCTGCTGCTG 1202
1884 GTAGAAACGGGATTTTACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1943
1203 GTAGAAACGGGATTTTACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
1944 CCACCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2003
1263 CCACCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
2004 AAGGGGAACTTTGAGAGAGGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCT 2063
1323 AAGGGGAACTTTGAGAGAGGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCT 1382
2064 GCACATGCTTATCTCTCCCATCTAGCCAGGATCTATGTTGTTTCTTCTG 2123
1383 GCACATGCTTATCTCTCCCATCTAGCCAGGATCTATGTTGTTTCTTCTG 1442
2124 TTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2183
1443 TTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
2184 GCAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2243
1503 GCAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1562
2244 GTTGTGTTCACT 2254
1563 GTTGTGTTCACT 1573

RESULT 10
BT007000 1146 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens putative protein similar to nesey (Drosophila) mRNA,
DEFINITION complete cds.
ACCESSION BT007000
VERSION BT007000.1 GI:30582838
KEYWORDS F11 CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1146)
AUTHORS Kainline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Phelan,M., and Farmer,A.
TITLE Cloning of human full-length cDNAs in BD Creator(TM) System Donor
vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1146)
AUTHORS Kainline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Phelan,M., and Farmer,A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA
COMMENT This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two

forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'Acc' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

```

1. 1146
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00356X1.0"
/clone_lib="BD Creator(TM) CDS library derived from MGC collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1. 1146
/codon_start=1
/product="putative protein similar to nesey (Drosophila)"
/protein_id="AAP35646.1"
/db_xref="GI:30582839"
/translation="MGRITTAVALTTRCFQMAVLLAGYVYTTANGNDIKTMPCVLTLLKILGLAVDFDGDQNSLSBQKVAIGVPSLLEAVSFYFGAFVGPFSNMHYMKLVOGLIDIPKIPNSIIPALKRLSLGLFYLVGYTLSPHTBEDYLDNHPFWRRCMYLIMGKFLVLYKVCWLVTEGVCILGLAGNGEKGKAKMDCAKMKWLFENPRMTGLASFNINAMVARIYFKRLKPLNGKLSGLIFLALHGHSGVLYCFQMKLIYIVERQARLQESPELSKLAITYLOPFYLVQOTIHMLEMGSMTAFCLTTWDMKLVVYSITPLGHIPLSLFPIPIYHKAMVPRKKLKME"

```

CDS

BASE COUNT 274 a 310 c 256 g 306 t

ORIGIN

Query Match 48.1%; Score 1094; DB 9; Length 1146;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1144; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 396 ATGGGCGGACCATCACTGCGCTCTCATCTACTCTTTGCTTCAGATGGCTACTCTTG 455
DB 1 ATGGGCGGACCATCACTGCGCTCTCATCTACTCTTTGCTTCAGATGGCTACTCTTG 60
QY 456 GCTGATTAATTACTGCTGCAACCGGCACTAGATATCAATGCAATGCCAATTGT 515
DB 61 GCTGATTAATTACTGCTGCAACCGGCACTAGATATCAATGCAATGCCAATTGT 120
QY 516 GTTCTGACTTTGAAGTGATGTTGGTGTGACTCTTGAACGAGGAGAAAGATCAG 575
DB 121 GTTCTGACTTTGAAGTGATGTTGGTGTGACTCTTGAACGAGGAGAAAGATCAG 180
QY 576 AATTCCTGTCTCTGAGCAACGAATAATGACATAGTGTGCTTCCCTGCTGGA 635
DB 181 AATTCCTGTCTCTGAGCAACGAATAATGACATAGTGTGCTTCCCTGCTGGA 240
QY 636 GTTGTGCTTTCTCTCACTTCTATGGGGCTTTCTTGGTAGGGCCCAAGTTTCATGAA 695
DB 241 GTTGTGCTTTCTCTCACTTCTATGGGGCTTTCTTGGTAGGGCCCAAGTTTCATGAA 300
QY 696 CACTACATGAAGTGTGAGGAGAGAGCTGATGACATCCAGAGAAAGATACCAACACG 755
DB 301 CACTACATGAAGTGTGAGGAGAGAGCTGATGACATCCAGAGAAAGATACCAACACG 360
QY 756 ATCAATTCCTGCTCTCAAGCGCTAGTCTGGGCTTTTCTTCACTAGTGGCTACACATG 815
DB 361 ATCAATTCCTGCTCTCAAGCGCTAGTCTGGGCTTTTCTTCACTAGTGGCTACACATG 420
QY 816 CTCAGGCCCCCATCAACAGAAAGCTATCTCTCACTAGAGACTATAGCAACACCCCTTC 875
DB 421 CTCAGGCCCCCATCAACAGAAAGCTATCTCTCACTAGAGACTATAGCAACACCCCTTC 480
QY 876 TGGTTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 935
DB 481 TGGTTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 936 TGGTTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 995

```

```

DB 541 TGGTTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 996 GAAAAGGCGAGGCAAGAGTGGATGCTTGTGCAACATGAGTGTGCTTTGAAATA 1055
DB 601 GAAAAGGCGAGGCAAGAGTGGATGCTTGTGCAACATGAGTGTGCTTTGAAATA 660
QY 1056 AACCCCGGCTTACAGGAGCACTATGCTCATTCATCAATCAACAGCAAGCCCTGGGAGCC 1115
DB 661 AACCCCGGCTTACAGGAGCACTATGCTCATTCATTCATCAATCAACAGCAAGCCCTGGGAGCC 720
QY 1116 CGCTACATCTTCAAGAGCACTCAAGTCTTGGAAATAAAGAACTCTTCAAGGCTCTCTG 1175
DB 721 CGCTACATCTTCAAGAGCACTCAAGTCTTGGAAATAAAGAACTCTTCAAGGCTCTCTG 780
QY 1176 TTGCTATTCCTGCGCTCTGAGCAAGGCTCTGCACTAGATATCTGTCTTCCAGATG 1235
DB 781 TTGCTATTCCTGCGCTCTGAGCAAGGCTCTGCACTAGATATCTGTCTTCCAGATG 840
QY 1236 GAATTCCTCATTTGTTATTTGGAAGAGAGAGGCTGAGGCTCATTCAGAGAGAGGAGCCAGCC 1295
DB 841 GAATTCCTCATTTGTTATTTGGAAGAGAGAGGCTGAGGCTCATTCAGAGAGAGGAGCCAGCC 900
QY 1296 CTGAGCAAGCTGCGCGCATTAATCTGCTCTCAGAGCCCTTCTACTATTTGGTGCAACAGACC 1355
DB 901 CTGAGCAAGCTGCGCGCATTAATCTGCTCTCAGAGCCCTTCTACTATTTGGTGCAACAGACC 960
QY 1356 ATCCACTGCTCTTCAATGAGTTACTGATGATGCTTCTGCTCTTCACTGAGGAGCAAA 1415
DB 961 ATCCACTGCTCTTCAATGAGTTACTGATGATGCTTCTGCTCTTCACTGAGGAGCAAA 1020
QY 1416 TGGCTTAAGTGTATTAATCCATCTATTCTTGGGACATCTTCTCTGAGGCTACTA 1475
DB 1021 TGGCTTAAGTGTATTAATCCATCTATTCTTGGGACATCTTCTCTGAGGCTACTA 1080
QY 1476 TTCAATTTGCTTAATTTCAACAAGCAATGATGCCAAGAGAAAGAGTTAAAGAGATG 1535
DB 1081 TTCAATTTGCTTAATTTCAACAAGCAATGATGCCAAGAGAAAGAGTTAAAGAGATG 1140
QY 1536 GAATA 1540
DB 1141 GAATA 1145

```

RESULT 11
BT007735 1146 bp mRNA linear SYN 13-MAY-2003
LOCUS
DEFINITION
Synthetic construct Homo sapiens putative protein similar to nesey (Drosophila) mRNA, partial cds.
ACCESSION
BT007735
VERSION
BT007735.1 GI:30584308
KEYWORDS
FLI-CDNA.
SOURCE
synthetic construct
ORGANISM
synthetic construct
REFERENCE
1 (bases 1 to 1146)
Kainline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Pheasant, M., and Farmer, A.
Cloning of human full-length cDNAs in BD Creator(TM) System Donor vector
TITLE
Unpublished
JOURNAL
2 (bases 1 to 1146)
Kainline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Pheasant, M., and Farmer, A.
Direct Submission
TITLE
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA
JOURNAL
This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal

tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

1. 1146
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH003561.0"
/clone_1lb="BD Creator(TM) CDS Library derived from MGC collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1. >1146
/note="Mutations: 1145:Stop->Leu"
/codon_start=1
/transl_table=1
/product="Homo sapiens putative protein similar to nesey (Drosophila)"
/protein_id="AAP36403.1"
/db_xref="GI:30584309"
/translation="MGRITAVLTTPGOMAYLLAGYVYATGNYDKMTMPHCVLT.
KLGLAVDFDGGKDONSLSSBOOKAIRQVPSLLEVAGSYRFGATLVGPQSMHY
MKLVQGLIDIPKIPNSIIIPALRSLSLGLFVLGYLLSLPHLEDYLLDNDHF
WFRGMVLMIGKFLYKYVTCMTLBEVCLTLGFGFGEKGAADAKMKNVLF
ETNRFEGTISFNINAWARYIPKRLFLSKLSEGLSLFLMLMGHLSGYLV
CPQKFLIVIVEROARLIOESPLSLKLAITVLOPYVLOQITHLFPYGSMTARC
LFTMDKMKLYKYSIVPLGHIFPLSLFILYHKHAWPRKXKMKML"

CDS

BASE COUNT 273 a 310 c 256 g 307 t
ORIGIN

Query Match 48.1%; Score 1093; DB 12; Length 1146;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 396 ATGGGCGGACCATCATCGCGTCTCACTACCTTTGGCTTCAAGATGGCTTACCTTCTG 455
DB 1 ATGGGCGGACCATCATCGCGTCTCACTACCTTTGGCTTCAAGATGGCTTACCTTCTG 60
QY 456 GCTGATATCTATTACATCGGACCGGCACTAGCATATCAAGTGGACATGGCCATTTGT 515
DB 61 GCTGATATCTATTACATCGGACCGGCACTAGCATATCAAGTGGACATGGCCATTTGT 120
QY 516 GTTCTGACTTGAAGTGAATGTTGGCTGTTGACTACTTGAAGGAGGAAAGATCAG 575
DB 121 GTTCTGACTTGAAGTGAATGTTGGCTGTTGACTACTTGAAGGAGGAAAGATCAG 180
QY 576 AATTCCTTGTCTCTGAGCAAGAAATATGCCATAGCTGTGTTCTTCTCTCTGAA 635
DB 181 AATTCCTTGTCTCTGAGCAAGAAATATGCCATAGCTGTGTTCTTCTCTCTGAA 240
QY 636 GTTCTGCTTCTCTCTCTCTCTCTATGGGCGCTTGTGGAGGCGCCAGTTCTCATGAT 695
DB 241 GTTCTGCTTCTCTCTCTCTCTCTATGGGCGCTTGTGGAGGCGCCAGTTCTCATGAT 300
QY 696 CACTACATGAAGTGTGACAGGAGAGTGAATACCAAGAAAGATACCAAGAGC 755
DB 301 CACTACATGAAGTGTGACAGGAGAGTGAATACCAAGAAAGATACCAAGAGC 360
QY 756 ATCATCTCTGCTCAAGCGCTGAGTCTGGGCTTTTCTAAGTGGGCTTACACTG 815
DB 361 ATCATCTCTGCTCAAGCGCTGAGTCTGGGCTTTTCTAAGTGGGCTTACACTG 420
QY 816 CTCGAGCCCGACATCAGAGAGACTATCTCTCACTGAGACTATGACCAACACCCCTTC 875
DB 421 CTCGAGCCCGACATCAGAGAGACTATCTCTCACTGAGACTATGACCAACACCCCTTC 480
QY 876 TGGTTCGCTGCATGATCATGCTGATCTGGGCAAGTTTGTCTGTAACAATATGTACCC 935
DB 481 TGGTTCGCTGCATGATCATGCTGATCTGGGCAAGTTTGTCTGTAACAATATGTACCC 540

QY 936 TGTGGCTGTGTACAGAGAGTATGATTTTGAACGGGCTTGAATGCTTTGAA 995
DB 541 TGTGGCTGTGTACAGAGAGTATGATTTTGAACGGGCTTGAATGCTTTGAA 600
QY 996 GAAAAGGCAAGGCAAGTGGAGTGTGCTGGCCATGAGAGTGTGCTTTGAACA 1055
DB 601 GAAAAGGCAAGGCAAGTGGAGTGTGCTGGCCATGAGAGTGTGCTTTGAACA 660
QY 1056 AACCCCGCTTCACTGAGCAAGTGGCTTCAATCAATCAACCAACGCTGGTGGCC 1115
DB 661 AACCCCGCTTCACTGAGCAAGTGGCTTCAATCAATCAACCAACGCTGGTGGCC 720
QY 1116 CGCTACATCTTCAAGACATGAGTCTTGGAAATGAAGAACTCTCAAGGCTCTCG 1175
DB 721 CGCTACATCTTCAAGACATGAGTCTTGGAAATGAAGAACTCTCAAGGCTCTCG 780
QY 1176 TTGCTATTCCTGAGCCCTTGGACAGGCTGACATCTGAGATACCTGCTTCCAGATG 1235
DB 781 TTGCTATTCCTGAGCCCTTGGACAGGCTGACATCTGAGATACCTGCTTCCAGATG 840
QY 1236 GAATTCCTCATTTGTTATTTGGAAGACAGGCTGCAAGGCTCATTAAGAGAGCCACCC 1295
DB 841 GAATTCCTCATTTGTTATTTGGAAGACAGGCTGCAAGGCTCATTAAGAGAGCCACCC 900
QY 1296 CTGAGCAAGTGGCCGCTTACTGCTTCCAGCCCTTCTACTATTGTTGCAACAGACC 1355
DB 901 CTGAGCAAGTGGCCGCTTACTGCTTCCAGCCCTTCTACTATTGTTGCAACAGACC 960
QY 1356 ATCCAGTGGCTTTCATGAGGTTACTGCATGAGTGCCTTGCCTTCAAGTGGGCAAA 1415
DB 961 ATCCAGTGGCTTTCATGAGGTTACTGCATGAGTGCCTTGCCTTCAAGTGGGCAAA 1020
QY 1416 TGGCTTAAGGTATTAATTCATCTATTTCTTGGCCACATCTTCTCTGAGCTTACTA 1475
DB 1021 TGGCTTAAGGTATTAATTCATCTATTTCTTGGCCACATCTTCTCTGAGCTTACTA 1080
QY 1476 TTGATATGCTTATATTCACAAAGCAATGCTGCCAAGGAAAGAAAGAAAGATG 1535
DB 1081 TTGATATGCTTATATTCACAAAGCAATGCTGCCAAGGAAAGAAAGAAAGATG 1140
QY 1536 GAAT 1539
DB 1141 GAAT 1144

RESULT 12
AC006512/c
LOCUS
DEFINITION Homo sapiens 12 PAC RP3-461P17 (Roswell Park Cancer Institute Human
PAC library) complete sequence.
ACCESSION AC006512
VERSION AC006512.13 GI:29469488
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 155975)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooke,S.L., Amaratunga,H.C., Aze,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dahorne,S.R., David,R.,
Davies,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Derr,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Hochs,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emelting,S.,
Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
Franz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,

GILL, R., Gorrell, J. H., Guevara, M., Gunaratne, P., Hale, S.,
 Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Haylak, P.,
 Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M.,
 Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S.,
 Hume, J., Ioshikhes, I., Jackson, L. E., Jacobson, B., Jia, Y.,
 Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U.,
 King, L., Korvan, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N.,
 Leal, B., Lee, E., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O.,
 Lien, C., Liu, J., Liu, W., Loussagad, H., Lozano, R. J., Lu, X.,
 Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Massey, P.,
 Marandel, I., Martin, R., Martindale, A., Martinez, E., Massey, E.,
 Mahoney, E., McLeod, M. P., Meador, M., Mei, G., Mescher, S.,
 Mezher, M., Miller, A., Miner, G., Miner, Z., Mitchell, T.,
 Mohabbat, K., Montgomery, K. T., Morgan, M., Morris, S., Moser, M.,
 Neill, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G.,
 Scherer, S., Scott, G., Shen, H., Shim, A., Shostkari, N., Sisson, I.,
 Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabori, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Umanai, K., Vasquez, L., Vera, V., Villalon, D., Visions, R., Wang, C.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
 Williams, G., Williamson, A., Wiczzyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Zuchelapacti, R.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (05-FEB-1999) Molecular and Human Genetics, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (18-JUN-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (02-JUN-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 5 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (03-JUN-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 6 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (18-JUN-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 7 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (127-OCT-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 8 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-MUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 9 (bases 1 to 155975)
 Worley, K.C.
 Direct Submission
 Submitted (129-M

```

TITLE      Direct Submission
JOURNAL    Submitted (02-APR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   10 (bases 1 to 155975)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (04-APR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Apr 2, 2003 this sequence version replaced gi:4926863.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
    STS are identified using ePCR (Genome Res. 7:541-550) searches
    of a local database that includes entries from dbSTS, GDB, and
    local mapping efforts.
    Repeats are identified using RepeatMasker (A. Smit and P. Green,
    unpublished.) for Human and Mouse sequences.
    Genes and Region of sequence similarity are identified by BLAST
    (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
    EST and cDNA sequences. Genes demonstrate at least two exons
    flanked by consensus splice sites that maintained sequence
    continuity across the splice junctions. Sequences that are not
    identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES
Source      .
            Location/Qualifiers
                1..155975
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="12"
                /clone="RP3-461F17"
                1..97210
                /note="overlaps bases 127585..222930 of clone"
                /function="clone overlap"
repeat_region
            1649..1772
                /rpt_family="L3b"
                complement(2781..3065)
repeat_region
            3127..3430
                /rpt_family="AluYd2"
                3127..3430
repeat_region
            3473..4048
                /rpt_family="AluSx"
                3473..4048
STS
repeat_region
            3592..3797
                /standard_name="GDB:214837"
                3592..3797
repeat_region
            4550..4771
                /rpt_family="MIR"
                4550..4771
repeat_region
            5127..5127
                /rpt_family="L3b"
                complement(4816..5127)
repeat_region
            5146..5288
                /rpt_family="AluSx"
                complement(5146..5288)

```


Db	18890	GTTGGCTGTGGGGGACGTGTGCAGAAACTACTCGTCCCTTTTCACAGCACTCCTTTG	18884
Oy	1613	CCCCAGACGAGAATTGAAAAACCAGGGAGGTGGAAGTCGATGCTTCAGCTGTGGCT	1672
Db	188840	CCCCAGACGAGAATTGAAAAACCAGGGAGGTGGAAGTCGATGCTTCAGCTGTGGCT	188781
Oy	1673	CTGCTGCCAGCCAGCTCTCATTTGGGGCCAAAGGGAAAATTCTTTTTTGGAGAGGCGT	1732
Db	188780	CTGCTGCCAGCCAGCTCTCATTTGGGGCCAAAGGGAAAATTCTTTTTTGGAGAGGCGT	188721
Oy	1733	CTTGTCTTGYCACCCACGCTGGAATGCAGTGGCGGATCTCAGCTCACCGCACTTCAC	1792
Db	188720	CTTGTCTTGTGTAACCAAGCTGGAATGCAGTGGCGGATCTCAGCTCACCGCACTTCAC	188651
Oy	1793	CTCTCGSGTTCAAAGTATTTTCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG	1852
Db	188660	CTCTCGSGTTCAAAGTATTTTCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG	188601
Oy	1853	CCACCATGCCGAGCTAATTTTGTATTTTCAGTAGAAAGGGATTCACACGTTGGCCA	1912
Db	188600	CCACCATGCCGAGCTAATTTTGTATTTTCAGTAGAAAGGGATTCACACGTTGGCCA	188541
Oy	1913	GGCTGTGTCGAAGTCTGTGACCGCAAGATCCACCGGCTCCGCTCCCAAGTGTGG	1972
Db	188540	GGCTGTGTCGAAGTCTGTGACCGCAAGATCCACCGGCTCCGCTCCCAAGTGTGG	188481
Oy	1973	GATTACAGGCGTGAGCCACCGTGCCCGGCCCAAGGGAAAATTCTTTGGGAGAGACGA	2032
Db	188480	GATTACAGGCGTGAGCCACCGTGCCCGGCCCAAGGGAAAATTCTTTGGGAGAGACGA	188421
Oy	2033	GGGGGTCAATCTCCCCTGTGATTCGCCCATGACAATGGCTTATCTCTCCCATCTAGC	2092
Db	188420	GGGGGTCAATCTCCCCTGTGATTCGCCCATGACAATGGCTTATCTCTCCCATCTAGC	188361
Oy	2093	CAGGATCTATGTGTTTTCTTCTGCCAATTACTATGATGTGATGTCGCGTACCA	2152
Db	188360	CAGGATCTATGTGTTTTCTTCTGCCAATTACTATGATGTGATGTCGCGTACCA	188301
Oy	2153	CCACCCCCCCTGAGGGGGGTGAGAGGGGTGACAGGCCCTGCTGCTCCACTTTTCTA	2212
Db	188300	CCACCCCCCCTGAGGGGGGTGAGAGGGGTGACAGGCCCTGCTGCTCCACTTTTCTA	188241
Oy	2213	CCTTGAAGTGTATTAAGATAAATCACTTCTGTTTGTTCAGTTTCTCA	2260
Db	188240	CCTTGAAGTGTATTAAGATAAATCACTTCTGTTTGTTCAGTTTCTCA	188193
RESULT 14			
LOCUS	BC009856/c	1641 bp	mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, Similar to gene rich cluster, C2f gene, clone		
ACCESSION	MGC:16358 IMAGE:3927564, mRNA, complete cds.		
VERSION	BC009856		
KEYWORDS	BC009856.1 GI:14602665		
SOURCE	MGC.		
ORGANISM	Homo sapiens (human)		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 1641)		
REFERENCE	Strausberg,R. Direct Submission Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgrabbs-r@mail.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LBNL) DNA Sequencing by: Institute for Systems Biology		

FEATURES	Source	Location/Qualifiers
CDs		<p>1. 1641</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="MGC:16358 IMAGE:3927564"</p> <p>/tissue_type="Ovary, adenocarcinoma"</p> <p>/clone_id="NIH_MGC_9"</p> <p>/lab_host="DH10B-R"</p> <p>/note="Vector: pOTB7"</p> <p>279..734</p> <p>/codon_start=1</p> <p>/product="Similar to gene rich cluster, C2f gene"</p> <p>/protein_id="AAH09856.1"</p> <p>/db_xref="GI:14602666"</p> <p>/transcript="MLMDSPLNRAAGLLOVITHOKNVLLEVNPOTRIPRDFSCGLM VOLHILKSVRADGPOKLEKIVINPVSDDHPVCMKVGTSFSLPVSVDRELVPSSDP IVTVVGAFAHGKRSVETTERKMSISNPLSALATCAKLTFAFEWNCVT"</p>
BASE COUNT	442 a 354 c 452 g 393 t	
ORIGIN		
Query Match	18.7% Score 424; DB 9; Length 1641;	
Best Local Similarity	100.0%; Pred. No. 1.9e-226; Mismatches 0; Indels 0; Gaps 0;	
Matches 424; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1837 TGGGATACAGGACGCGCACCATGCGCAGCTAATTTTGTATTTTCAGTAGAAACGGGAT	1896
DB	1533 TGGGATATACAGGACGCGCACCATGCGCAGCTAATTTTGTATTTTCAGTAGAAACGGGAT	1474
QY	1897 TTCAACCAAGTTGGCCAGGCTGGTTCGAATCTCTGACCGCAAGTATCCACCCGCTCCG	1956
DB	1473 TTCAACCAAGTTGGCCAGGCTGGTTCGAATCTCTGACCGCAAGTATCCACCCGCTCCG	1414
QY	1957 CCTCCCAAAGTGTGGGATTTACAGGCGGTGAGCCACGCTGCCCGGCCCAAAGGGGAAACTC	2016
DB	1413 CCTCCCAAAGTGTGGGATTTACAGGCGGTGAGCCACGCTGCCCGGCCCAAAGGGGAAACTC	1354
QY	2017 TTGTGGAGGAGCAGAGGGGCTCATCTCCCTCTGATTTCCCATGACATTTGCTTA	2076
DB	1353 TTGTGGAGGAGCAGAGGGGCTCATCTCCCTCTGATTTCCCATGACATTTGCTTA	1294
QY	2077 TCTCTCCCATCTAGCCAGGATTTATTTGTTTCTTCTGCCAATTTACTATGATTTGT	2136
DB	1293 TCTCTCCCATCTAGCCAGGATTTATTTGTTTCTTCTGCCAATTTACTATGATTTGT	1234
QY	2137 GTATGTGCGGTACACACACCCCCCATAGGGGGGGGTGAGAGGGGGTGCAAGGCCCTGCC	2196
DB	1233 GTATGTGCGGTACACACACCCCCCATAGGGGGGGGTGAGAGGGGGTGCAAGGCCCTGCC	1174
QY	2197 TGCTCCACTTTTCTTACTTTGGAAGTATATATTAATACATTTCTGTTTGTTCAGTTT	2256
DB	1173 TGCTCCACTTTTCTTACTTTGGAAGTATATATTAATACATTTCTGTTTGTTCAGTTT	1114
QY	2257 TTCA	
DB	1113 TTCA	
RESULT 15		
BD023986	BD023986	450 bp DNA linear PAT 27-AUG-2002
LOCUS	BD023986	450 bp DNA linear PAT 27-AUG-2002
DEFINITION	Sequence tag and encoded human protein.	
ACCESSION	BD023986.1	GI:22565209
VERSION	BD023986.1	GI:22565209
KEYWORDS	JP 2001269182-A/232.	

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 232 02-OCT-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/232
PD 02-OCT-2001
PR 24-FEB-2000 JP 2000118773
PI 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC score 3.8
CC seq VVALAGVLSGFQ/EL
FH key Location/Qualifiers
FT CDS 67..450
FT sig_peptide 67..138.
FT Location/Qualifiers
FEATURES
source 1..450
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 71 a 130 c 122 g 125 t 2 others
ORIGIN
Query Match 17.7%; Score 402; DB 6; Length 450;
Best Local Similarity 100.0%; Pred. No. 4.6e-214;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 GGGGGGTCCCTGTGGGGCTCCCGAGTTAAGATGCGCTCTCAGCGGAGGGAGCGAGG 106
DB 36 GGGGGGTCCCTGTGGGGCTCCCGAGTTAAGATGCGCTCTCAGCGGAGGGAGCGAGG 95
QY 107 GACTGTGTGTGGCGCGTGGGGGTTCTGCAATCGGGTTTCCAGAGCTGAGCCTTAACA 166
DB 96 GACTGTGTGTGGCGCGTGGGGGTTCTGCAATCGGGTTTCCAGAGCTGAGCCTTAACA 155
QY 167 GTTGGGAGCGTCCCTGGGCGCGTCAAGAACAGGCGCTGCGCTGATCATCTCCATCTTCT 226
DB 156 GTTGGGAGCGTCCCTGGGCGCGTCAAGAACAGGCGCTGCGCTGATCATCTCCATCTTCT 215
QY 227 GGGTTACCCCTTGTGTTGTTTATGCGCATTAACCTTTTACAGAGAGACTACCTCAT 286
DB 216 GGGTTACCCCTTGTGTTGTTTATGCGCATTAACCTTTTACAGAGAGACTACCTCAT 275
QY 287 CCACTCTTCCATACCTTTACAGAGCGCTCAATGCTTATTTTAACCTTTGGAACCAAGCT 346
DB 276 CCACTCTTCCATACCTTTACAGAGCGCTCAATGCTTATTTTAACCTTTGGAACCAAGCT 335
QY 347 CTACCACTCCCTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 406
DB 336 CTACCACTCCCTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 395
QY 407 CATCACTGCGCTCACTACCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 448
DB 396 CATCACTGCGCTCACTACCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 437

Search completed: November 22, 2003, 10:33:30
Job time : 8293 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:20:19 ; Search time 40 Seconds
(without alignments)
3141.790 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 487
Sequence: 1 MASSAEGDEGTVALAGVLQ.....ILPYHKAMVPRKELKKME 487

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	381	78.2	381	4 Q92980	Q92980 homo sapien
2	280	57.5	381	4 Q9BW40	Q9BW40 homo sapien
3	50	10.3	374	11 Q8BNH6	Q8BNH6 mus musculu
4	50	10.3	440	11 Q51311	Q51311 mus musculu
5	50	10.3	487	11 Q91V01	Q91V01 mus musculu
6	50	1.8	458	5 Q961A6	Q961A6 drosophila
7	9	1.8	515	5 Q81PE9	Q81PE9 drosophila
8	9	1.8	556	5 Q9VL88	Q9VL88 drosophila
9	9	1.8	1462	10 Q9FIC5	Q9FIC5 arabidopsis
10	9	1.8	1486	10 Q9FXB7	Q9FXB7 arabidopsis
11	8	1.6	78	6 Q9GUX2	Q9GUX2 sus scrofa
12	8	1.6	86	6 Q8MN94	Q8MN94 oryctolagus
13	8	1.6	87	6 Q9TSG2	Q9TSG2 sus scrofa
14	8	1.6	87	13 Q8QFT5	Q8QFT5 gallus galli
15	8	1.6	299	10 Q8W0Y7	Q8W0Y7 medicago tr
16	8	1.6	541	17 Q97BR4	Q97BR4 thermoplasm

17	8	1.6	547	16 Q92288	Q92288 rhizobium m
18	8	1.6	588	10 Q9FIC0	Q9FIC0 arabidopsis
19	8	1.6	599	10 Q93VM9	Q93VM9 arabidopsis
20	8	1.6	599	10 Q8H134	Q8H134 arabidopsis
21	8	1.6	824	2 Q8VNZ4	Q8VNZ4 bordetella
22	8	1.6	1233	4 Q9UK61	Q9UK61 homo sapien
23	8	1.6	1854	5 Q9VTY8	Q9VTY8 drosophila
24	8	1.4	13	11 Q62354	Q62354 mus musculu
25	7	1.4	37	11 Q9OV33	Q9OV33 rattus ap.
26	7	1.4	54	16 Q92G83	Q92G83 rickettsia
27	7	1.4	55	16 Q97D05	Q97D05 streptococ
28	7	1.4	67	16 Q8NU02	Q8NU02 streptococ
29	7	1.4	82	10 Q93XT1	Q93XT1 diospyros h
30	7	1.4	91	2 Q9ACD7	Q9ACD7 vibrio salm
31	7	1.4	93	15 Q79599	Q79599 human immun
32	7	1.4	100	16 Q8CYQ8	Q8CYQ8 streptococ
33	7	1.4	101	3 Q07236	Q07236 saccharomyc
34	7	1.4	102	4 Q9B210	Q9B210 homo sapien
35	7	1.4	103	2 Q8RPN0	Q8RPN0 entlichia c
36	7	1.4	110	12 Q8VAV8	Q8VAV8 white spot
37	7	1.4	115	4 Q9UDH5	Q9UDH5 homo sapien
38	7	1.4	118	16 Q8F6V0	Q8F6V0 leptospira
39	7	1.4	129	17 Q8U237	Q8U237 pyrococcus
40	7	1.4	132	10 Q9SQ73	Q9SQ73 goseypium h
41	7	1.4	132	10 Q9SQ74	Q9SQ74 goseypium h
42	7	1.4	143	16 Q8ED53	Q8ED53 shewanella
43	7	1.4	150	17 Q9HSP8	Q9HSP8 halobacteri
44	7	1.4	154	4 Q9Y628	Q9Y628 homo sapien
45	7	1.4	157	2 Q8RR49	Q8RR49 streptococ

ALIGNMENTS

RESULT 1

Q92980 ID Q92980 PRELIMINARY; PRT; 381 AA.
AC Q92980;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE C3F.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Ansari-Lari M.A., Shen Y., Muzny D.M., Lee W., Gibbs R.A.,
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Ansari-Lari M.A., Shen Y., Muzny D.M., Lee W., Gibbs R.A.,
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U72515; AAC51640.1; -.

DR EMBL: U47924; AAB51326.1; -
 DR InterPro: IPR001064; Crystallin.
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT_1.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
 SQ SEQUENCE 381 AA; 4412 MW; FDNFA658155113 CRC64;

Query Match 78.2%; Score 381; DB 4; Length 381;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 MGRITAVITFCFQMAVYLLAGYYTATGNVDIKMTMPCVLTLLKLGAVDPGCKQ 166
 DB 1 MGRITAVITFCFQMAVYLLAGYYTATGNVDIKMTMPCVLTLLKLGAVDPGCKQ 60
 QY 167 NSLSSEQOKYAIRGVPSLLEVAGFSYFGAPLVGQPSMNHYMKLVQSELIDIPKIPNS 226
 DB 61 NSLSSEQOKYAIRGVPSLLEVAGFSYFGAPLVGQPSMNHYMKLVQSELIDIPKIPNS 120
 QY 227 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNHFWRCMYMLIWGKFLYKYVT 286
 DB 121 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNHFWRCMYMLIWGKFLYKYVT 180
 QY 287 CWLVTEGVCILTGLGFNGFEKQKAKMDACANMKVWLFETNPRFTGTIASFNINNAWA 346
 DB 181 CWLVTEGVCILTGLGFNGFEKQKAKMDACANMKVWLFETNPRFTGTIASFNINNAWA 240
 QY 347 RYIFRKLFLGNKELSGSLSLFLALMHGSHGYLVCFMELIYVVEQARLQESPT 406
 DB 241 RYIFRKLFLGNKELSGSLSLFLALMHGSHGYLVCFMELIYVVEQARLQESPT 300
 QY 407 LSKLAITVLOPFYLVQTIHMLFWGYSMTAFCLFTMDKMLKVYSIYFLGHIFELSL 466
 DB 301 LSKLAITVLOPFYLVQTIHMLFWGYSMTAFCLFTMDKMLKVYSIYFLGHIFELSL 360
 QY 467 FILPYIHKAMVPRKEKLKME 487
 DB 361 FILPYIHKAMVPRKEKLKME 381

RESULT 2

Q9BW40 PRELIMINARY; PRT; 381 AA.
 AC Q9BW40;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Strusberg R.;
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC000664; AA000664.1; -
 DR InterPro: IPR001064; Crystallin.
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT_1.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 381 AA; 4412 MW; 2BF19AB389F39B1 CRC64;

Query Match 57.5%; Score 280; DB 4; Length 381;
 Best Local Similarity 100.0%; Pred. No. 1.3e-284;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 MGRITAVITFCFQMAVYLLAGYYTATGNVDIKMTMPCVLTLLKLGAVDPGCKQ 166
 DB 1 MGRITAVITFCFQMAVYLLAGYYTATGNVDIKMTMPCVLTLLKLGAVDPGCKQ 60

QY 167 NSLSSEQOKYAIRGVPSLLEVAGFSYFGAPLVGQPSMNHYMKLVQSELIDIPKIPNS 226
 DB 61 NSLSSEQOKYAIRGVPSLLEVAGFSYFGAPLVGQPSMNHYMKLVQSELIDIPKIPNS 120
 QY 227 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNHFWRCMYMLIWGKFLYKYVT 286
 DB 121 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNHFWRCMYMLIWGKFLYKYVT 180
 QY 287 CWLVTEGVCILTGLGFNGFEKQKAKMDACANMKVWLFETNPRFTGTIASFNINNAWA 346
 DB 181 CWLVTEGVCILTGLGFNGFEKQKAKMDACANMKVWLFETNPRFTGTIASFNINNAWA 240
 QY 347 RYIFRKLFLGNKELSGSLSLFLALMHGSHGYLVCFM 386
 DB 241 RYIFRKLFLGNKELSGSLSLFLALMHGSHGYLVCFM 280

RESULT 3

O8BNH6 PRELIMINARY; PRT; 374 AA.
 AC O8BNH6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Unknown Est (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Body;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK083687; BAC38993.1; -
 FT NON_TER 1
 SQ SEQUENCE 374 AA; 43259 MW; 68E4BD6B72A98486 CRC64;

Query Match 10.3%; Score 50; DB 11; Length 374;
 Best Local Similarity 100.0%; Pred. No. 2.5e-43;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 GTIASFNINNAWARYIFKRLKFLGNKELSGSLSLFLALMHGSHGYL 381
 DB 219 GTIASFNINNAWARYIFKRLKFLGNKELSGSLSLFLALMHGSHGYL 268

RESULT 4

O35131 PRELIMINARY; PRT; 440 AA.
 AC O35131;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE C3f (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96112780; PubMed=9445485;
 RA Ansari-Lari M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
 RA Lu J., Gorrell J.H., Chinault A.C., Belmont J.W., Miller W.,
 RA Gibbs R.A.;
 RT "Comparative sequence analysis of a gene-rich cluster at human
 RT chromosome 12p13 and its syntenic region in mouse chromosome 6.";
 RL Genome Res. 8:29-40(1998).
 RN [2]

RP SEQUENCE FROM N.A.
RA Murthy D., Ansaari-Lari M.A., Timms K.M., Yu W., Dugan S., Lu J.,
RA Shen Y., Rowland K., Liu W., Perez L., Ding Y., Haywood M., Jain A.,
RA Leal B., Logan O., Nguyen V., Savage L., Shen H., Morley K., Chen E.,
RA Forcum J., Arenson A.D., Chiu M.W., Correll J.H., Brundage E., Di W.,
RA Chihault C., Nelson D., Gibbs R.A.;
RT "Direct Submission";
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Muzny D., Arenson A.D., Adams C., Brundage E., Bunac C., Carvelli K.,
RA Chacko J., Chen J., Di W., Ding Y., Dugan S., Durbin J., Forcum J.,
RA Ganesh R., Garcia C., Goodman M., Gorrell J.H., Haywood M.,
RA Hernandez J., Jackson L., Jin S., Kamal R., Karpachy S., Kovar C.,
RA Leal B., Li Y., Lichtarge O., Liu W., Logan O., Lu J., Ly T.,
RA Martinez C., Osval G., Perez L., Rashid N.D., Rowland K., Savage L.,
RA Scherer S.E., Shen H., Simon M., Stovall K., Timms K.M., Todd J.,
RA Vo Q., Williamson A., Worley K.C., Yu W., Chihault C., Nelson D.,
RA Gibbs R.A.;
RT "Direct Submission";
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC002397; AAC36007.1; -
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT; 1.
DR NON_TER
SQ SEQUENCE 440 AA; 5117 MW; 7A78BAAB40A1C638 CRC64;

Query Match 10.3%; Score 50; DB 11; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.9e-43; Indels 0;
Matches 50; Conservative 0; Mismatches 0; Gaps 0;

QY 332 GTTASFNNTNAAWARYIFKRLKFLGNKELSGSLFLALMHLGHSYL 381
Db 285 GTTASFNNTNAAWARYIFKRLKFLGNKELSGSLFLALMHLGHSYL 334

RESULT 5
Q91V01 PRELIMINARY; PRT; 487 AA.
ID Q91V01
AC Q91V01;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative transmembrane protein PTG (Unknown) (Protein for
MGC:11670).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RA Zhu Y., Han Y., Reddy J.K.;
RT "Cloning and initial characterization of mouse PTG cDNA, whose
expression is in a PPAR alpha dependent manner";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY028317; AAK02915.1; -
DR EMBL: BC006753; AA06753.2; -
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT; 1.
DR Transmembrane.
SQ SEQUENCE 487 AA; 56146 MW; EBC51DB3734B17C7 CRC64;

Query Match 10.3%; Score 50; DB 11; Length 487;
Best Local Similarity 100.0%; Pred. No. 3.1e-43; Indels 0;
Matches 50; Conservative 0; Mismatches 0; Gaps 0;

QY 332 GTTASFNNTNAAWARYIFKRLKFLGNKELSGSLFLALMHLGHSYL 381

Db 332 GTTASFNNTNAAWARYIFKRLKFLGNKELSGSLFLALMHLGHSYL 381

RESULT 6
Q961A6 PRELIMINARY; PRT; 458 AA.
ID Q961A6
AC Q961A6;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE LD25907P (CG3779-PC).
GN Drosophila melanogaster (Fruit fly).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frie E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Biazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Bortman M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin J.K., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jaitai M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Kerchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,
RA Williams S.M., Woodgett M., Wray K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,

Qy 360 ELSQGLSL 368
 Db 328 ELSQGLSL 336

RESULT 8
 Q9VLB8 PRELIMINARY; PRT; 556 AA.

AC Q9VLB8; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE NUMB protein.
 GN NUMB OR CG3779.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champs M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotter P.,
 RA Butts K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gairdian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,
 RA Jaitani M., Kaulish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkova R., Tector C., Turner K., Ventler E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Weissman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195 (2000).
 DR EMBL; AE003625; AAF5276.1; -
 DR HSSP; P16554; 2NMB.
 DR FLYBase; FBgn0002973; numB.
 DR InterPro; IPR006020; PTB_PTD.
 DR Pfam; PF00640; PTD_1.
 DR SMART; SM00462; PTB; 1.
 DR PROSITE; PS01179; PTD; 1.
 SQ SEQUENCE 556 AA; 60614 MW; 91B2695B5DB8405 CRC64;

Query Match 1.8%; Score 9; DB 5; Length 556;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 360 ELSQGLSL 368
 Db 369 ELSQGLSL 377

RESULT 9
 Q9FXB7 PRELIMINARY; PRT; 1462 AA.

AC Q9FXB7; 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Retroelement pol polyprotein-like.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucosida II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=99156233; PubMed=10048488;
 RA Asanizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones";
 RL DNA Res. 5:379-391 (1998).
 DR EMBL; AB016892; BAB10837.1; -
 DR EMBL; AB010694; BAB10837.1; JOINED.
 DR InterPro; IPR005162; Retrotrans_gag.
 DR Pfam; PF03732; Retrotrans_gag; 1.
 DR Pfam; PF00665; rve; 1.
 SQ SEQUENCE 1462 AA; 163504 MW; DA9700D7BFCB900B CRC64;

Query Match 1.8%; Score 9; DB 10; Length 1462;
 Best Local Similarity 100.0%; Pred. No. 8.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 SIFUGYPPA 55
 Db 716 SIFUGYPPA 724

RESULT 10
 Q9FXB7 PRELIMINARY; PRT; 1486 AA.

AC Q9FXB7; 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative retroelement polyprotein.
 GN F25P12.89.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucosida II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=99156233; PubMed=10048488;
 RA Asanizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones";
 RL DNA Res. 5:379-391 (1998).
 DR EMBL; AB016892; BAB10837.1; -
 DR EMBL; AB010694; BAB10837.1; JOINED.
 DR InterPro; IPR005162; Retrotrans_gag.
 DR Pfam; PF03732; Retrotrans_gag; 1.
 DR Pfam; PF00665; rve; 1.
 SQ SEQUENCE 1462 AA; 163504 MW; DA9700D7BFCB900B CRC64;

Query Match 1.8%; Score 9; DB 10; Length 1462;
 Best Local Similarity 100.0%; Pred. No. 8.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF03732; Retrotrans_gag; 1.
 DR Pfam; PF06665; rve; 1.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS50158; ZF_CCHC; 1.
 KW PolyProtein.
 SQ SEQUENCE 1466 AA; 165725 MW; CAF70FA04CEBA690 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 1486;
 Best Local Similarity 100.0%; Pred. No. 8.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 SIFLGYPFA 55
 |||||
 Db 752 SIFLGYPFA 760

RESULT 11

Q9GJX2 PRELIMINARY; PRT; 78 AA.
 AC Q9GJX2, 01-MAR-2001 (TRENBLREL. 16, Created)
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Diazepam binding inhibitor (Fragment).
 GN DBI.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Skeletal muscle;
 RA Davoli R., Fontanesi L., Zambonelli P., Bigi D., Milc J., Cagnazzo M.,
 RA Gencl V., Braglia S., Russo V.;
 RT "Isolation and mapping of expressed sequence tags from porcine
 RT skeletal muscle: a contribution to the genomic transcript map of this
 RT tissue."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ301367; CAC21172.1; -.
 DR HSSP; P07107; IACA.
 DR InterPro; IPR000582; Ac_coa_bind_prot.
 DR Pfam; PF00887; ACP; 1.
 DR PRINTS; PR00689; ACOABINDINGP.
 DR ProDom; PD002965; Ac_coa_bind_prot; 1.
 DR PROSITE; PS00880; ACP; 1.
 FT NON TER 78
 RN [1]
 SQ SEQUENCE 78 AA; 8807 MW; 55D1861C52EC6FAC CRC64;

Query Match 1.6%; Score 8; DB 6; Length 78;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 308 KKGAKWDA 315
 |||||
 Db 51 KKGAKWDA 58

RESULT 12

Q8WN94 PRELIMINARY; PRT; 86 AA.
 AC Q8WN94, 01-MAR-2002 (TRENBLREL. 20, Created)
 DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Acyl-CoA-binding protein (ACBP) (Diazepam binding inhibitor) (DBI)
 DE (Endozepine) (EP).
 GN DBI.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Barnack N.H., Liu H., Qian Z., Bilderback T.;
 RT "Localization of an endogenous diazepam ligand, acyl CoA binding
 RT protein, to Muller cells in rabbit retina."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: BINDS MEDIUM- AND LONG-CHAIN ACYL-COA ESTERS WITH VERY
 CC HIGH AFFINITY AND MAY FUNCTION AS AN INTRACELLULAR CARRIER OF
 CC ACYL-COA ESTERS. IT IS ALSO ABLE TO DISPLACE DIAZEPAM FROM THE
 CC BENZODIAZEPINE (BZD) RECOGNITION SITE LOCATED ON THE GABA TYPE A
 CC RECEPTOR. IT IS THEREFORE POSSIBLE THAT THIS PROTEIN ALSO ACTS AS
 CC A NEUROPEPTIDE TO MODULATE THE ACTION OF THE GABA RECEPTOR (BY
 CC SIMILARITY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ACP FAMILY.
 CC EMBL; AF407578; AAK98608.2; -.
 DR InterPro; IPR000582; Ac_coa_bind_prot.
 DR Pfam; PF00887; ACP; 1.
 DR PRINTS; PR00689; ACOABINDINGP.
 DR ProDom; PD002965; Ac_coa_bind_prot; 1.
 DR PROSITE; PS00880; ACP; FALSE_NEG.
 KW Transport; Lipid-binding; Acetylation.
 FT INIT MET 0
 FT MOD RES 1 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 86 AA; 9784 MW; F55EF76BB434A0F1 CRC64;

Query Match 1.6%; Score 8; DB 6; Length 86;
 Best Local Similarity 100.0%; Pred. No. 7.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 308 KKGAKWDA 315
 |||||
 Db 50 KKGAKWDA 57

RESULT 13

Q9TSG2 PRELIMINARY; PRT; 87 AA.
 AC Q9TSG2, 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Endozepine.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Back Fat;
 RA Rauli N., Suzuki H., Karnuah A.B., Hamastina N.;
 RT "Cloning of pig endozepine gene."
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB019792; BAA34531.1; -.
 DR HSSP; P07107; IACA.
 DR InterPro; IPR000582; Ac_coa_bind_prot.
 DR Pfam; PF00887; ACP; 1.
 DR PRINTS; PR00689; ACOABINDINGP.
 DR ProDom; PD002965; Ac_coa_bind_prot; 1.
 DR PROSITE; PS00880; ACP; 1.
 SQ SEQUENCE 87 AA; 9896 MW; 3225B53E8307DE38 CRC64;

Query Match 1.6%; Score 8; DB 6; Length 87;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 308 KKGAKWDA 315
 |||||
 Db 51 KKGAKWDA 58

RESULT 14

Search completed: November 21, 2003, 14:24:00
 Job time : 42 secs

Q8QFT5 PRELIMINARY; PRT; 87 AA.
 AC Q8QFT5;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Diazepam binding inhibitor.
 GN DBI;
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=broiler breeder; TISSUE=liver;
 RA Baader M.;
 RT "Drug-induction of Hepatic Cytochromes P450: Studies in Chicken
 Hepatoma Cells.";
 RL Thesis (2002), Department of Biozentrum Pharmacology/Neurobiology,
 University of Basel, Basel, Switzerland.
 DR EMBL; AJ430511; CAD33129.1; -;
 DR InterPro; IPR000582; AC_coa_bind_prot. 1.
 DR Pfam; PF00887; ACBP; 1.
 DR ProDom; PD002965; AC_coa_bind_prot; 1.
 SQ SEQUENCE 87 AA; 9805 MW; 67405B1C62AB066 CRC64;

Query Match 1.6%; Score 8; DB 13; Length 87;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGKAKMDA 315
 |||||
 Db 51 KGKAKMDA 58

RESULT 15
 Q8W0Y7 PRELIMINARY; PRT; 299 AA.
 AC Q8W0Y7;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Enod8.3 (Fragment).
 GN ENOD8.3.
 OS Medicago truncatula (Barrel medic).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 OX NCBI_TaxID=3880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Jemalong;
 RA Dickstein R., Hu X., Yang J., Ba L., Coque L., Kim D.-J., Cook D.R.,
 RA Yeung A.T.;
 RT "Randomly duplicated Enod8 genes are differentially expressed in
 Medicago.";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF463407; AAL68830.1; -;
 DR InterPro; IPR01087; Lipase_GDSL.
 DR Pfam; PF00657; Lipase_GDSL; 1.
 FT NON_TER 1
 SQ SEQUENCE 299 AA; 33075 MW; F2A7D59A499471AB CRC64;

Query Match 1.6%; Score 8; DB 10; Length 299;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 KIPNSIIP 229
 |||||
 Db 33 KIPNSIIP 40

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:11:33 / Search time 39 Seconds
(without alignments)
3222.349 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 2580
Sequence: 1 MASSAEGDEGTVALAGVLAQ.....ILPYIKRAMPYRKELKKME 487

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2314	89.7	487	11 Q91V01	Q91V01 mus musculu
2	2174	84.3	440	11 Q35131	Q35131 mus musculu
3	2046	79.3	381	4 Q92980	Q92980 homo sapien
4	2042	79.1	381	4 Q9BM40	Q9BM40 homo sapien
5	1856	71.9	374	11 Q8BNH6	Q8BNH6 mus musculu
6	874.5	33.9	497	5 Q9YVY9	Q9YVY9 drosophila
7	871.5	33.8	497	5 Q9VXVS	Q9VXVS drosophila
8	766	29.7	473	5 Q9V125	Q9V125 caenorhabdi
9	408.5	15.8	722	5 Q9V5F2	Q9V5F2 drosophila
10	407	15.8	465	3 Q9CAN8	Q9CAN8 arabidopsis
11	406	15.7	509	3 Q42916	Q42916 schizosacch
12	406	15.7	619	3 Q85448	Q85448 saccharomyc
13	383.5	14.9	462	10 Q8RWT4	Q8RWT4 arabidopsis
14	383.5	14.9	480	5 Q18856	Q18856 caenorhabdi
15	382.5	14.8	457	10 Q9LNB3	Q9LNB3 arabidopsis
16	373.5	14.5	519	11 Q8R312	Q8R312 mus musculu

17	373	14.5	492	11 Q8BRH8	Q8BRH8 mus musculu
18	372	14.4	492	11 Q8R3T2	Q8R3T2 mus musculu
19	344.5	13.4	487	11 Q9C273	Q9C273 mus musculu
20	341.5	13.2	487	11 Q8R192	Q8R192 mus musculu
21	300	11.6	386	11 Q8BRH5	Q8BRH5 mus musculu
22	297	11.5	540	5 Q17826	Q17826 caenorhabdi
23	280	10.9	378	11 Q8BWS6	Q8BWS6 mus musculu
24	266	10.3	350	4 Q96KY4	Q96KY4 homo sapien
25	263.5	10.2	451	5 Q8T802	Q8T802 drosophila
26	260.5	10.1	473	11 Q9CT76	Q9CT76 mus musculu
27	260	10.1	350	3 Q03130	Q03130 saccharomyc
28	255.5	9.9	473	11 Q8R1P9	Q8R1P9 mus musculu
29	254.5	9.9	472	4 Q96N66	Q96N66 homo sapien
30	254.5	9.9	473	11 Q8CHK3	Q8CHK3 mus musculu
31	244.5	9.5	480	5 Q19468	Q19468 caenorhabdi
32	242	9.4	303	4 Q8N9R5	Q8N9R5 homo sapien
33	241.5	9.4	489	5 Q9VMD5	Q9VMD5 drosophila
34	213.5	8.3	450	13 Q91938	Q91938 xenopus lae
35	212	8.2	343	4 Q9BPV2	Q9BPV2 homo sapien
36	212	8.2	455	13 Q91936	Q91936 xenopus lae
37	210.5	8.2	456	13 Q91937	Q91937 xenopus lae
38	208	8.1	461	13 Q91935	Q91935 xenopus lae
39	190	7.4	450	4 Q9H234	Q9H234 homo sapien
40	188.5	7.3	344	4 Q9BRE9	Q9BRE9 homo sapien
41	185	7.2	450	11 Q9J3J9	Q9J3J9 mus musculu
42	185	7.2	456	11 Q9J3J9	Q9J3J9 mus musculu
43	184	7.1	456	4 Q9H236	Q9H236 homo sapien
44	183.5	7.1	455	11 Q9J3J8	Q9J3J8 mus musculu
45	180.5	7.0	461	11 Q9J3J7	Q9J3J7 mus musculu

ALIGNMENTS

RESULT 1

ID	Q91V01	PRELIMINARY;	PRT;	487 AA.
AC	Q91V01;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Putative transmembrane protein P7G (Unknown) (Protein for MGC:11670).			
DE	Mus musculus (Mouse).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RC	STRAIN=C57BL/6J; TISSUE=Liver;			
RT	Zhu Y., Han Y., Reddy J.K.;			
RT	"Cloning and initial characterization of mouse P7G cDNA, whose expression is in a PPAR alpha dependent manner."			
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N. A.			
RA	Strauberg R.;			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY028317; AK20915.1; -			
DR	EMBL; BC006753; AA06753.2; -			
DR	InterPro; IPR004299; MBOAT_fam.			
DR	Pfam; PF03062; MBOAT; 1.			
KW	Transmembrane.			
SQ	SEQUENCE 487 AA; 56146 MW; EBC51DB3734B17C7 CRC64;			

Query Match	89.7%; Score 2314; DB 11; Length 487;
Best Local Similarity	87.3%; Pred. No. 1,1e-188;
Matches	425; Conservative 30; Mismatches 32; Indels 0; Gaps 0;
Qy	1 MASSAEGDEGTVALAGVLAQSGFOELSLNKLATSLGASEQALRIITIFLCYPPALFYRH 60
Db	1 MASTADDMGRTLEQMRGLWPGVDELINKLATSLGASEQALRIITIFLCYPPALFYRH 60

```

QY 61 YLFYKETYLIHFFHTFTGSIAYFNGNOLYHSLCTIVQFLIRLMGRITAVLTTCF 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 YLFYKDSYLIHFFHTFTGSIAYFNGHGFPHSLCVALQFLIRLMGRITAVLTTCF 120
QY 121 QMAVYLLAGYATGATGNYDIKMTMPHCVTLLKLGAVDYFDGKONSLSSEQKAIIRG 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 QMAVYLLAGYATGATGNYDIKMTMPHCVTLLKLGICIDYDGDGKONSLSSEQKAIIRG 180
QY 181 VESLLEVAAGSYFYGAFLVGPFSNMHYMKLVQGLIDIPGKIPNSTIPALRLSLGLFY 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 VESLLEVAAGSYFYGAFLVGPFSNMHYMKLVQGLTIDIPGKIPNSTIPALRLSLGLFY 240
QY 241 LVGYTLSPHITEDYLLTEDYDNNHPFRCMMLIGKFLVYKXYTCMLVTEGVCILTL 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 LVGYTLSPHITEDYLLTEDYDNNHPFRCMMLIGKFLVYKXYTCMLVTEGVCILTL 300
QY 301 GNGEPEEKAKMDACANMKVLFETNPRFTGTIASFNINTNMAVARIYFKELKFLGNKE 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 GNGEPEENGTVWDACANMKVLFETNPRFTGTIASFNINTNMAVARIYFKELKFLGNKE 360
QY 361 LSQGLSLFLALMHGHSGLVLCQFQMEFLIVYERQAARLIOESPITSLAITVLOPEY 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 LSQGLSLFLALMHGHSGLVLCQFQMEFLIVYERQAARLIOESPITSLAITVLOPEY 420
QY 421 YLVQQTIIHMLFMGYSTAFCLFTWDKMLKYKSIYFLGHIFFLSLFLPIYIHKAMVPRK 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 421 YLVQQTIIHMLFMGYSTAFCLFTWDKMLKYKSIYFLGHIFFLSLFLPIYIHKAMVPRK 480
QY 481 EKLKKME 487
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 481 EKLKKRE 487

```

RESULT 2

```

ID 035131 PRELIMINARY; PRT; 440 AA.
AC 035131;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE C3F (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112780; PubMed=9445485;
RA Ansari-Lari M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
RA Lu J., Gorrell J.H., Chinault A.C., Belmont J.W., Miller W.,
RA Gibbs R.A.;
RT "Comparative sequence analysis of a gene-rich cluster at human
RT chromosome 12p13 and its syntenic region in mouse chromosome 6.";
RL Genome Res. 8:29-40(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Shen Y., Ansari-Lari M.A., Timms K.M., Yu W., Dugan S., Lu J.,
RA Shen Y., Rowland K., Liu W., Perez L., Ding Y., Haywood M., Jain A.,
RA Leal B., Logan O., Nguyen V., Savage L., Shen H., Worley K., Chen E.,
RA Forcum J., Arenson A.D., Chiu M.W., Gorrell J.H., Brundage E., Di W.,
RA Chinault C., Nelson D., Gibbs R.A.;
RT "Direct Submission.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Muzny D., Arenson A.D., Adams C., Brundage E., Bunac C., Carvelli K.,
RA Chacko J., Chen J., Di W., Ding Y., Dugan S., Durbin J., Forcum J.,
RA Ganesh R., Garcia C., Goodman M., Gorrell J.H., Haywood M.,
RA Hernandez J., Jackson L., Jin S., Kampel R., Karpathy S., Kovar C.,
RA Leal B., Li Y., Lichtege O., Liu W., Logan O., Lu J., Ly T.,
RA Martinez C., Owal G., Perez L., Rashid N.D., Rowland K., Savage L.,
RA Scherer S.E., Shen H., Simon M., Stovall K., Timms K.M., Todd J.,
RA Vo O., Williamson A., Worley K.C., Yu W., Chinault C., Nelson D.,

```

```

RA Gibbs R.A.;
RT "Direct Submission.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002397; AAC36007.1; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT.1.
FT NON TER
SQ SEQUENCE 440 AA; 5117 MW; 7A78BA840A1C638 CRC64;

```

```

Query Match 84.3%; Score 2174; DB 11; Length 440;
Best Local Similarity 89.7%; Pred. No. 8.2e-177;
Matches 393; Conservative 25; Mismatches 20; Indels 0; Gaps 0;

```

```

QY 50 LGYPALPYRHYLFPEKETYLIHFFHTFTGSIAYFNGNOLYHSLCTIVQFLIRLMGR 109
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 LGYPALPYRHYLFPEKDSYLIHFFHTFTGSIAYFNGHGFPHSLCVALQFLIRLMGR 62
QY 110 TITAVLTTCFQMAVYLLAGYATGATGNYDIKMTMPHCVTLLKLGAVDYFDGKONSLS 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 63 TITAVLTTCFQMAVYLLAGYATGATGNYDIKMTMPHCVTLLKLGICIDYDGDGKONSLS 122
QY 170 SSEQKAIIRGYPSLLEVAAGSYFYGAFLVGPFSNMHYMKLVQGLIDIPGKIPNSTIP 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 123 TSEQKAIIRGYPSLLEVAAGSYFYGAFLVGPFSNMHYMKLVQGLTIDIPGKIPNSTIP 182
QY 220 ALKRLSLGLFYLVGYTLSPHITEDYLLTEDYDNNHPFRCMMLIGKFLVYKXYTCML 289
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 183 ALKRLSLGLVYLVGYTLSPHITEDYLLTEDYDNNHPFRCMMLIGKFLVYKXYTCML 242
QY 290 VTEGVCILTLGFGNFEKAKMDACANMKVLFETNPRFTGTIASFNINTNMAVARIY 349
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 243 VTEGVCILTLGFGNFDENGTVWDACANMKVLFETNPRFTGTIASFNINTNMAVARIY 302
QY 350 FPKLFLGNKELSGSLSLFLALMHGHSGLVLCQFQMEFLIVYERQAARLIOESPITSLK 409
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 303 FPKLFLGNKELSGSLSLFLALMHGHSGLVLCQFQMEFLIVYERQAARLIOESPITSLK 362
QY 410 LAAITVLOPEFYLVQQTIIHMLFMGYSTAFCLFTWDKMLKYKSIYFLGHIFFLSLFL 469
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 363 LAAITVLOPEFYLVQQTIIHMLFMGYSTAFCLFTWDKMLKYKSIYFLGHIFFLSLFL 422
QY 470 PYIHKAMVPRKELKME 487
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 423 PYIHKAMVPRKELKRE 440

```

RESULT 3

```

ID 092980 PRELIMINARY; PRT; 381 AA.
AC 092980;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE C3F.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ansari-Lari M.A., Shen Y., Muzny D.M., Lee W., Gibbs R.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96303695; PubMed=8723724;
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
RA Malley T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
RT genes at human chromosome 12p13.";
RL Genome Res. 6:314-326(1996).
RN [3]
RP SEQUENCE FROM N.A.

```

RA Anseri-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
 RA Malley T., Gibbs R.A.;
 RA Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Anseri-Lari M.A., Shen Y., Muzny D.M., Lee W., Gibbs R.A.;
 RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Anseri-Lari M.A., Shen Y., Muzny D.M., Lee W., Gibbs R.A.;
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U72515; AAC51640.1; -
 DR EMBL: U47924; AAB51326.1; -
 DR InterPro: IPR001064; Crystalin.
 DR InterPro: IPR004239; MBOAT_fam.
 DR Pfam: PF03062; MBOAT_1.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
 SQ SEQUENCE 381 AA; 44122 MW; FD2AF658155113 CRC64;

Query Match 79.3%; Score 2046; DB 4; Length 381;
 Best Local Similarity 100.0%; Pred. No. 5.5e-166;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 MGRITAVLTTCFQMAVLLAGYTTATGNNDIKWTMPCVTLTKLIGLAVDPFGKQD 166
 DB 1 MGRITAVLTTCFQMAVLLAGYTTATGNNDIKWTMPCVTLTKLIGLAVDPFGKQD 60
 QY 167 NSLSEQOKYAIRGVPSLLEVAGFSYFGAPLVGPQFSNMHYMKLVQGLIDIPKIPNS 226
 DB 61 NSLSEQOKYAIRGVPSLLEVAGFSYFGAPLVGPQFSNMHYMKLVQGLIDIPKIPNS 120
 QY 227 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNPFWRCMYMLIWGKFLVKYVT 286
 DB 121 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNPFWRCMYMLIWGKFLVKYVT 180
 QY 287 CMLVTEGVCILTGFGNGPEEKAKMDACANMKWLFETNRPFTGTIASFNINTNAWA 346
 DB 181 CMLVTEGVCILTGFGNGPEEKAKMDACANMKWLFETNRPFTGTIASFNINTNAWA 240
 QY 347 RYIFRKLFLGNKELSGSLFLALMHGSHGYLVCFQMEFLIYVEROARLIQESPT 406
 DB 241 RYIFRKLFLGNKELSGSLFLALMHGSHGYLVCFQMEFLIYVEROARLIQESPT 300
 QY 407 LSKLAITVLOPFYLVQOTIHWLFMGYSMTAFCLFTWDKMLVKYSIYFLGHIFLSL 466
 DB 301 LSKLAITVLOPFYLVQOTIHWLFMGYSMTAFCLFTWDKMLVKYSIYFLGHIFLSL 360
 QY 467 FILPYIHKAMVPRKEXLKKME 487
 DB 361 FILPYIHKAMVPRKEXLKKME 381

RESULT 4

Q9BNH6 PRELIMINARY; PRT; 381 AA.
 ID Q9BNH6
 AC Q9BNH6
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Stiusberg R.;
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC000664; AAH00664.1; -
 DR InterPro: IPR001064; Crystalin.
 DR InterPro: IPR004239; MBOAT_fam.
 DR Pfam: PF03062; MBOAT_1.

DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 381 AA; 44121 MW; 2BF19AB369FF39B1 CRC64;

Query Match 79.1%; Score 2042; DB 4; Length 381;
 Best Local Similarity 99.7%; Pred. No. 1.2e-165;
 Matches 380; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 107 MGRITAVLTTCFQMAVLLAGYTTATGNNDIKWTMPCVTLTKLIGLAVDPFGKQD 166
 DB 1 MGRITAVLTTCFQMAVLLAGYTTATGNNDIKWTMPCVTLTKLIGLAVDPFGKQD 60
 QY 167 NSLSEQOKYAIRGVPSLLEVAGFSYFGAPLVGPQFSNMHYMKLVQGLIDIPKIPNS 226
 DB 61 NSLSEQOKYAIRGVPSLLEVAGFSYFGAPLVGPQFSNMHYMKLVQGLIDIPKIPNS 120
 QY 227 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNPFWRCMYMLIWGKFLVKYVT 286
 DB 121 IIPALKRSLGLFYLVGYTLLSPHITEDYLLTEDYDNPFWRCMYMLIWGKFLVKYVT 180
 QY 287 CMLVTEGVCILTGFGNGPEEKAKMDACANMKWLFETNRPFTGTIASFNINTNAWA 346
 DB 181 CMLVTEGVCILTGFGNGPEEKAKMDACANMKWLFETNRPFTGTIASFNINTNAWA 240
 QY 347 RYIFRKLFLGNKELSGSLFLALMHGSHGYLVCFQMEFLIYVEROARLIQESPT 406
 DB 241 RYIFRKLFLGNKELSGSLFLALMHGSHGYLVCFQMEFLIYVEROARLIQESPT 300
 QY 407 LSKLAITVLOPFYLVQOTIHWLFMGYSMTAFCLFTWDKMLVKYSIYFLGHIFLSL 466
 DB 301 LSKLAITVLOPFYLVQOTIHWLFMGYSMTAFCLFTWDKMLVKYSIYFLGHIFLSL 360
 QY 467 FILPYIHKAMVPRKEXLKKME 487
 DB 361 FILPYIHKAMVPRKEXLKKME 381

RESULT 5

Q9BNH6 PRELIMINARY; PRT; 374 AA.
 ID Q9BNH6
 AC Q9BNH6
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Unknown EST (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Body;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RL Nature 420:565-573 (2002)."
 DR EMBL: AK083687; BAC38993.1; -
 FT NON_TER 1
 SQ SEQUENCE 374 AA; 43259 MW; 68E4BD6B72A98486 CRC64;

Query Match 71.9%; Score 1856; DB 11; Length 374;
 Best Local Similarity 89.8%; Pred. No. 8.1e-150;
 Matches 336; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 114 VLTTCFQMAVLLAGYTTATGNNDIKWTMPCVTLTKLIGLAVDPFGKQDQNSLSEQ 173
 DB 1 VLTTCFQMAVLLAGYTTATGNNDIKWTMPCVTLTKLIGLAVDPFGKQDQNSLSEQ 60
 QY 174 OKYAIRGVPSLLEVAGFSYFGAPLVGPQFSNMHYMKLVQGLIDIPKIPNSIIPALKR 233
 DB 61 OKYAIRGVPSLLEVAGFSYFGAPLVGPQFSNMHYMKLVQGLIDIPKIPNSIIPALKR 120

```

QY 234 LSLGLFYLGVYTLSPHITTEDEYLLTDYDNDHPFPCOMWLIMGKFLVLYKYVTGMLVTBG 293
DB 121 LSLGLFYLGVYTLSPHITTEDEYLLTDYDNDHPFPCOMWLIMGKFLVLYKYVTGMLVTBG 180
QY 294 VCILITGLFNGFEKAKADACANNKWLFEETNPFTGTIASFNINTANAVARIYFKRL 353
DB 181 VCILITGLFNGFEKAKADACANNKWLFEETNPFTGTIASFNINTANAVARIYFKRL 240
QY 354 KFLGNKELSGSLTLFLALMHGSHGYVCFQMEFLIVVERQAAALIQESPTLSKLAAL 413
DB 241 KFLGNKELSGSLTLFLALMHGSHGYVCFQMEFLIVVERQAAALIQESPTLSKLAAL 300
QY 414 TVLQPFYLVQOQTHMLFMGYSMTAFCLFTWMDKLKYKSIYFLGHIFLSTLFLPYTH 473
DB 301 TALQPFYLVQOQTHMLFMGYSMTAFCLFTWMDKLKYKSIYFLGHIFLSTLFLPYTH 360
QY 474 KAMVPRKELKKME 487
DB 361 KAMVPRKELKKME 374

```

RESULT 6

```

QYXV9 PRELIMINARY; PRT; 497 AA.
AC 09XV9;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative transmembrane protein NES5.
DE NES OR NES5 OR CG9655.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
OX NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RA Maurel-Zaffran C., Chauvet S., Miasod R., Jullien N., Pradel J.,
RA Aragnol D.,
RT "The evolutionary conserved gene nes5 is a putative target of Hox
RT proteins in Drosophila embryo."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF28112; AAD28257.1; -
DR FlyBase: FBgn0026630; nes-
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT; 1.
KW Transmembrane.
SQ SEQUENCE 497 AA; 57436 MW; FBF535B13700520B CRC64;

```

```

Query Match 33.9%; Score 874.5; DB 5; Length 497;
Best Local Similarity 37.2%; Pred. No. 4.4e-66;
Matches 167; Conservative 92; Mismatches 165; Indels 25; Gaps 5;
QY 28 LNKLTSLGASQALRLISIFLPYRHYL-FYKETYLIHLFTFTGLSIAYFNR 86
DB 14 MDGIASGVGVPAALRLILLAGYPAALYQKFIIVAKYVHNHFACGAGLCYFNR 73
QY 87 GNOLYSLCTV-LQFLILRLKRTTAVLTTCPCOMAVYLLAGYATGATNDIKMTMP 145
DB 74 GRDTYSLILTLTYFLVLLLRKKTQFLAINFVFNHSYLLCYFYSSNDYDILMTMP 133
QY 146 CVLTLLKLGAVYFPGGKQNSLSEQKATIRGVPSLLEVAGSFYAFVGPQFSM 205
DB 134 CILVLRNIGYGFITDGLKESELSKQKTAALKRPSELELAFTFSGFLVGFQFPR 193
QY 206 NHYMKLVQGEILIDIPKIPNSIIPALKSLSLGLFYLGVYTLSPHITTEDEYLLTDYDNDHP 265
DB 194 RRRKAFVDFGRHQEGVEAGV----RRFGAGFYLLVQVGYGRYLPDSYFLTPPEFAQVS 249
QY 266 FWRRCMYMLIWGFVLVLYKYVTGMLVTGVCILITGLFNGFEKAKADACANNKWLFE 325
DB 250 FVRIYLLGFWAKFSYKYISWLLTEGALICIGLTFYKGBDKNGQPDWMSGCSNVKLTLE 309

```

```

QY 326 TNPRTGTIASFNINTANAVARIYFKRLKFLGNKELSGSLTLFLALMHGSHGYVCFQ 385
DB 310 TGMTNHHVQSPNNTNMGVGYITKRLKFLNNRITISGALGFLAVHGHYSGYMTL 369
QY 386 MEFLIVVERQAAAL-----IQESPTLSKLAALTVLQPFYLVQOQTHMLFMGY 434
DB 370 MEYVAVSTEKQTRFYTQVVLQPMGHILNNSDIYKL-----LYFITLKSYNVVMGM 421
QY 435 SMTAFCLFTWMDKLKYKSIYFLGHIFL 463
DB 422 CLTAFLVFLKERVIVYAGVSYGGFTFLV 450

```

RESULT 7

```

QYXV5 PRELIMINARY; PRT; 497 AA.
AC 09VX5;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE NES protein (R203440P).
DE NES OR CG9655.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
OX NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RA MEDLINE=20199606; PubMed=10731132;
RC STRAIN=BERKELEY;
RX ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodali C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mialshina N.V., Mobarry C., Morris J., Moshirefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy B., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN=berkeley;

```

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.,
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO03517; AAF49181.1; -
DR EMBL: AY070936; AAL48558.1; -
DR FlyBase: FBgn0026630; nes
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT; 1.
SQ SEQUENCE 497 AA; 57393 MW; 4927EFBD784BF861 CRC64;

Query Match 33.8%; Score 871.5; DB 5; Length 497;
Best Local Similarity 37.2%; Pred. No. 7.8e-66;
Matches 167; Conservative 92; Mismatches 165; Indels 25; Gaps 5;

QY 28 LNKLTATSLGASEQALRLIISIFLGYPPALFYRHYL-FYKETYLIHLFPTTGLSIAVFNQ 86
DB 14 MDGIASGVVPEALRLITLITLILAGYPAALYQKISIVADKTVHMFAGGAGLCTFNY 73
QY 87 GNOLYHSLCTV-LQFLILRMGRTTAVLTTCFQMAVLLAGYYTATGNVDIKMTMPH 145
DB 74 GLDTYHSLILITTYFLVLLRKKTQIFLAINFVFNHMSYLLGYFTSSNDYDILMTMPH 133
QY 146 CVLTLKLGAVDPDGDKNQSLSSBOQKAIKRGVSLLEAVGSFYFGAFVINGPOFSM 205
DB 134 CILVLRMIGFDTTDDKESELSKQKPKPSLELLAFSPSGELVGPQPP 193
QY 206 NHYMKLVQGLIDIPKIPNSIIPALKRLSLGLFVLYGYTLLSPHITEDYLLTEDYDNH 265
DB 194 RRYAFADGFERQHEGVNAGV---RRFGAFYLLVCQGLAYRLDSYFLTEPFAQS 249
QY 266 FWFECMTMLWGKVLKYVTCTLVTSVCITLGLGFNGFEKKAQKADCAKMKWLF 325
DB 250 FVKRIYLLGFMAKPSLYKXISCMILTGGALICILYKGBDKNQCPMWSGCSNKLLE 309
QY 326 TNPFTGISFNTNMAVARYFKRLKPLGNKELSGSLFLALMHGHSGLVCFQ 385
DB 310 TGNTEHYVQSFNTQWQOYIKRKLKFNKNTISGALLGLAWHGHSGLVYTF 369
QY 386 MEFLIVERQAPRL-----IQESPILSKLAITVLOPFYVVOQTIMHLMFGY 434
DB 370 MEWVWVTEKQITFYRKVVLPQMGHILNMSDIKL-----LYFITLKSANVVWGMW 421
QY 435 SMTAFCLFTMDKMLKVTKSYFLGHIFPL 463
DB 422 CLTAFVFLKXERMIWVGAVSYGFTFLV 450

RESULT 8
001925 PRELIMINARY; PRT; 473 AA.
AC 001925;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein R155.1.
GN R155.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genomic sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gaisel C., Mameley P., Kramer J.;
RT "The sequence of C. elegans coenid R155.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003390; AAB54271.2; -
KM Hypothetical protein.
SQ SEQUENCE 473 AA; 54113 MW; 2198AA79247D1A77 CRC64;

Query Match 29.7%; Score 766; DB 5; Length 473;
Best Local Similarity 33.8%; Pred. No. 7.1e-57;
Matches 157; Conservative 95; Mismatches 203; Indels 10; Gaps 3;

QY 31 LATSLSGSEQALRLIISIFLGYPPALFYRHYL-FYKETYLIHLFPTTGLSIAVFNQ 90
DB 7 LSEVTSASEDALRLISVLAGYPAALYHRTFYNNKRAQHQHLEFVIGSLMFMFGSSV 66
QY 91 YHSLCTVLOFLILRMGRTTAVLTTCFQMAVLLAGYYTATGNVDIKMTMPH 150
DB 67 IHPILSIFGAFYTNFMAQTDASIVAHIVFLGHLILGYWFHEHETDYDITMTTPTFCIMTL 126
QY 151 KLIGLAVDPDGDKNQSLSSBOQKAIKRGVSLLEAVGSFYFGAFVINGPOFSM 210
DB 127 RFLGLVMDVVDGQKQKPEHLKPDQKLTISDKPGLLETAAGLFFQGLTVLPQFLSKFRS 186
QY 211 LVQGLIDIPKIPNS-IIPALKRLSLGLFVLYGYTLLSPHITEDYLLTEDYDNH 269
DB 187 FVNGDMILDSGQPKPKAFPLPSISGFLAGCTYMWLHQGQWIDQYNSDANNLSFMR 246
QY 270 CMTMLWGKVLKYVTCTLVTSVCITLGLGFNGFEKKAQKADCAKMKWLFETNPR 329
DB 247 MSWVTLFRLTWKXCYCAMWILTEGASISGLIGHNGKABENDWDGVRDHIIMKMGHD 306
QY 330 FTGTGISFNTNMAVARYFKRLKPLGNKELSGSLFLALMHGHSGLVCFQ 389
DB 307 YNSVESFNGCTVPAFANHRRIRWNNKLASHVITLSYLAHGHGSLFLGVELG 366
QY 390 IVIVERQAPRLIQESPILSKLAITVLOPFYVVOQTIMHLMFGYSTAFCLFTMDKMLK 449
DB 367 CVAQNLQVLLIKRTPMSAISKPSRPIWIFGKLTISYSGAFALMGLITKXWIG 426
QY 450 VVKSIFYLGHIF-----LSLFTLP-YIHKAMPKREKLLK 485
DB 427 PVKSLYFIFGIYFVWPIILHMLRLVLPKPKAAAEKPEVKK 471

RESULT 9
09V5F2 PRELIMINARY; PRT; 722 AA.
AC 09V5F2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CG18445 protein.
GN CG18445.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherter S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.P., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palzozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Messarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*";
 RA Science 287:2185-2195(2000);
 DR EMBL: AE003831; AAF58858.1; -
 DR Flybase: FBgn003476; CG18445.
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT; 1.
 SQ SEQUENCE 722 AA; 81221 MW; CBAE84PA6720E367 CRC64;

Query Match 15.8%; Score 408.5; DB 5; Length 722;
 Best Local Similarity 25.8%; Pred. No. 3.3e-26;
 Matches 121; Conservative 89; Mismatches 212; Indels 47; Gaps 14;
 31 LATSICASQALRLISIFLGYPAFLFYHYFYKE--LYLLHLFTFTGSIAYFNFN 88
 22 LADWGLSLVDVNFILCOISALFLASLFRSMHPSVSKLHRTPLSLGIAFGFCFQ 81
 89 QLYH----SLLCIVLQFLILRLMGRTTAVLTTF--CFOMAYLLAGYYTATGNDIKW 141
 82 QAIHLAGLPAICIVIRTOPRIQPAVLVAMSYLLCVHLMRQLYDY----GSYALDI 136
 142 TMRPHCVLTKLIGLVANDYFDG--GKQNSLSSQOKAIGVSLLEVAQSFYGAFLVG 200
 137 TGPLMTITQKVSGLAFSIHDFVGRGDELTAKQOAHAIKMPDALEFSGYVHNPSTILAG 196
 201 POFSMNHMYKLVQ--ELIDIP---GKIPNS-----IIP-----ALKRLSG---LFYLVG 243
 197 PLVFKDYIEFVPEGNYLLSTPPGNGVLDSSKREVLPEPTPAVIRKVGSLVCAFIEMK 256
 244 YTLASP--HITEDYLLTEDYDHPFRFCMYLIMGKFLVLYKXYTCWLTGVCILTGIG 301
 257 FPKIYVVKMKEDDDFNNTSMYKYV---YAMMATTCIRFKYHAWMLADAI CNNSGIG 312
 302 ENGFEKGAQKADACANMKWLFETNPRTGTIASNINTNMAVARIKRLKFLGNKL 361
 313 FPGYDGDGSKWDLSINIVLSFEESTNRDANNNNCSTNRRLRTLYVERVP---QQY 368
 362 SQGLSLFLALMHGSLGVLVCFQMFELIVIERQAARLIOSEPSTLSKLAATVLOPFY 421
 369 GILTLTALAAVHMGFPGYLLTFAITCAVAVTAAKTRRLFRH----RQOSTVTRMFD 423
 422 LVQQTILHLMFGVSMATFCLFTWDMKLKYYKSIYFLGHIFLSLFLP 470

DB 424 ILTCITRVLVGATFPFVLLEFMGSIKLYRFLYLCIHLITSLVTFILP 472
 RESULT 10
 ID 09CAN8 PRELIMINARY; PRT; 465 AA.
 AC 09CAN8;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN F16M19.14 OR AT1G63050.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eustrosia II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cnv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etmu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldenrich A.D., Haas B., Hansen N.P., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maltl R., Marziani A.,
 RA Miltseher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Southwick A.M.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Uiterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RA "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*
 RA thaliana";
 RT Nature 408:816-820(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banb J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Ondaera C.S., Quach H.L.,
 RA Tang C., Tortum M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers A.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene At1g63050 (GI15221667)".;
 RT Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.M.,
 RA Lee J.M., Quach H.L., Tang C., Tortum M., Wu H.C., Yu G., Bowser L.,
 RA Carrinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Arabidopsis Open Reading Frame (ORF) Clones".;
 RT Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC010795; AA051612.1; -
 DR EMBL: AY072080; AAL59903.1; -
 DR EMBL: AY122979; AAM67512.1; -
 DR InterPro: IPR001064; Cysteal1in.
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT; 1.
 DR PROSITE: PS00225; CRYSTALLIN_BETACAMMA; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 465 AA; 52769 MW; 9A35E91B8277DCB3 CRC64;

Query Match 15.8%; Score 407; DB 10; Length 465;
 Best Local Similarity 25.4%; Pred. No. 2.6e-26;
 Matches 123; Conservative 90; Mismatches 227; Indels 44; Gaps 12;

QY 26 LSLNKLATSLGASEOALRLIISIFLGYPALFRRHYLFYKETYLIHLPHFTGSLIAFVN 85
 DB 4 LDMNSMASIGVSVLRFLLCFVATIPISFLMR--FLPSRLGHIYSAASGALSYLS 60
 QY 86 FGNOLYHSLCTIVLOFLRLMGRITTA-----LTFPCFQMAVLGAYYATANG-- 137
 DB 61 FG-----FSSNLHFLVPMPTIGVASMATYRPLSGPITFFLGPAVLICGHVFMGDAWK 113
 QY 138 --DIKMPHCYLTLLKLGAVDFDGDONSLSSEQKXAIRGVPSLLEVAQSYFYG 195
 DB 114 EGGIDSTGALMLVTLKVISCSINVDGMLKEGSLAEAKKRLIOMPSLIEFQCLCCG 173
 QY 196 AFLVGPQSMNHMYKLVOGE--LIDIPGKIPNSIIPALKRISLGLFYLVGYTLLSPHT 252
 DB 174 SHFAPRVPEMKDYLETWTEKGIWAVESEKGRPSYGAAMI RAVQALCMALYLVLPQP 233
 QY 253 EDYLLTEDYDNRHPEFRCMYMLIMGKFVLYKYVTGMLVTEGVCILTGFGNF--EEKK 310
 DB 234 LTRFTEPYOEWKFLKRFQYQWAGFTARKYFTIWSISEASTIISGLFGSGWDETQTK 293
 QY 311 AKMDACAMKWLFEFTNPREFTGTIASFNINTNAAVARYIFKRLKFLGNK-----LSQG 364
 DB 294 AKMDAKKVDILGLVLAASAQIPLFPMI QVSTMLRHYERIVRPGKAGFOLLATQT 353
 QY 365 LSLFLALHGLHSGYLVCFQMEFLIYVERQAARLIQ--SEPTLSKLAITVLOPFYLV 423
 DB 354 VS---AVMHGLYPCYIIFVFQALMIDSKAIYRMOQALPPKAMLRNVLLINFLYLV 409
 QY 424 QOTIHWLMGYSMTAFCLFTMDKMLKYKSIYFLGHIFLSLFLPIHKAIVRKRKL 483
 DB 410 -----VVLNYSVSGFVNLSEHETLVAFKSVYIGTIVPIAVL--LLSYL--VPVKVRPRT 461

QY 484 KME 487
 DB 462 RKE 465

RESULT 11
 Q42916 PRELIMINARY; PRT; 509 AA.

AC 042916;
 DT 01-JAN-1999 (TREMBLrel. 09, Created)
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN SPBC16A3.10.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972n-;
 RL Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R,
 RA Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL021748; CA16861.1; -
 DR Genedb Spombe; SPBC16A3.10; -
 DR InterPro; IPR004299; MBOAT_fam.
 DR Pfam; PF03062; MBOAT; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 509 AA; 58821 MW; 74B4E3E3D9DE65D9 CRC64;

Query Match 15.7%; Score 406; DB 3; Length 509;
 Best Local Similarity 25.1%; Pred. No. 3.5e-26;
 Matches 121; Conservative 86; Mismatches 206; Indels 70; Gaps 13;

QY 25 ELINKLATSLGASEOALRLIISIFLGYPALFRRHY-----LFYKETYLI--HLPHFTGSLI 69

DB 6 DIPEYFSSFLGVHPDQLKLLFCFLSAVFPFGILKRLPSAPWIRNLFSSISGLF-----YL 61
 QY 70 IHLPHFTGSLIAFVN-----FGNOLYHSLCTIVLOFLRLMGRITTAVLTTFCQWAX 124
 DB 62 IGHHLLVDGVLVLLFDLFTYFAAFRRSSHMPWIIFFV--ILGHTSSHVIRYIY---- 115
 QY 125 LLAGYVYATANGYDIKMTMPHCYLTLLKLGAVDFDGDONSLSSEQKXAIRGVPSL 184
 DB 116 -----PSNTIPI--TASQWUCLMKLTARMSVYDRLSSLSSTYQKRALRKIPNI 165
 QY 185 LEVAGFSYCAFLVGPQSMNHMYKLVO---GELIDI PK--IPNSIIPALKRISLGL 238
 DB 166 LYFLGVYFPPSLVGLVGFADYDERFTLSMFKPLADPYEKQTPHSLBGRGCRGL 225
 QY 239 FYLVGYTLLSHIETDYLITTEDYDNRHPEFRCMYMLIMGKFVLYKYVTGMLVTEGVCILT 298
 DB 226 LMLLFTGSSIVPLKFLTPKFASSPILLKRYGVCITAFARMKYYGAMELSDACILS 285
 QY 299 GLGFNGFEKAKMDACAMKWLFEFTNPREFTGTIASFNINTNAAVARYIFKRLKFLGN 358
 DB 286 GIGINGLDSGHPMDRKNIDPIKFEFADNIKALEMANNNTKMLRNYVLAVAKGK 345
 QY 359 KE--LSQSLILFL--ALMHGLHSGYLVCFQMEFLIYVERQAARLIQ-----ESPTLS 408
 DB 346 RPFKSTLSTFTVSAAMHGVAGYLLFVSAFIQTVAKYRRHVRPFLKPMDET--- 402
 QY 409 KLAITVLOPF---YLVQOTIHWLMFGYSMTAFCLFTMDKMLKYKSIYFLGHIFLSL 465
 DB 403 -----GPKRYVDYGVATNLSLYLISFLNLKESIHVMKELYPIVHIYILIA 454
 QY 466 LFI 468
 DB 455 LAV 457

RESULT 12
 Q08548 PRELIMINARY; PRT; 619 AA.

AC 008548;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Chromosome XV reading frame ORF YOR175C.
 GN YOR175C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE OF 269-619 FROM N.A.
 RA Bordonne R., Camases A., Madania A., Martin R.P., Poch O.,
 RA Tarasov I.A., Winzor B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hughes B., Pohl T.M.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MIRS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 275083; CA99384.1; -
 DR SGD; S0005701; YOR175C.
 DR InterPro; IPR004299; MBOAT_fam.
 DR Pfam; PF03062; MBOAT; 1.
 SQ SEQUENCE 619 AA; 72227 MW; 87C7AC9194BE0BC5 CRC64;

Query Match 15.7%; Score 406; DB 3; Length 619;
 Best Local Similarity 26.2%; Pred. No. 4.5e-26;
 Matches 133; Conservative 79; Mismatches 238; Indels 58; Gaps 17;

QY 28 LNKLATSLGASEOALRLIISIFLGYPALFRRHY-----LFYKETYLI--HLPHFTGSLI 81

```

Db      9 LKRLITNYGIDSEFTLEAYICLGSFPLNALIKRIPEKRIGLKCCFIISMSEFYL----- 63
Qy      82 AIFENQNOYHSLCLICVLOFLIRLMGRITITAVLTFFCQOMAYLAGYYTATGNDIKR 141
Db      64 GVALNVSGRFTLPSTMTFTYLLSRFYRSKEMPHL-NFMFVGHILAIINHIAOFLMEQOT 122
Qy      142 TM-----PHCVLTLLKLGIAVDYFDG-----KDQNSLSEQOKAIRGVSLLEVAQSYF 193
Db      123 TYDISSQNVLMAMKLTSFAMSYDGSCTSESPKDLTEHOKRAVGHGHPPLKFLAYAF 182
Qy      194 YAAFLVGPQFSNMHYMKLVQGEI-DIP-----GKIPNSIIPALKRSLGL 238
Db      183 YSTLLTGPSFDYADPFSMNLCEMFRLPESSKKPMRRHNGERQIPKNGKALMKVQGL 242
Qy      239 FVLVGTLLSPHITETLYLTED-YDNHPFWRFCMYMLINGKFLYKYVTCMLTBEVCVL 297
Db      243 AMWILSTLGMKHPVRYVLDKQGFPRSEIFRIHYLFLGFIHRFYAAMTISEGSCYL 302
Qy      298 TGLGFNGFEER-CKAKMDACAMKVMLEPTNPRFTGSIASFNINNAVARYIFRKLKFL 356
Db      303 CELGVNGVSKOKIRMDVRNIDITVETAQNTREMLAMNNTKMLKYSVYLAVTK 362
Qy      357 GKKEISQGLSLFL-ALMHGLSGYLVCFQMEFLIVEROAARLIQES--PTSKLAA 412
Db      363 GKPPGFRSTLFTFLTSAPMGHTRPGYLLFATGATLY-----QTCGKIYRNNRPIFLRBDG 418
Qy      413 ITVL--QPFYIVQOQTIHMLFMQYSMTAFCLFTWDMKLVKYSIYLGHIIF--FLSLFI 468
Db      419 VPLPSKRLYDLVGIYAIKARFVQVPIILDLKPSLWWSVYVYVHIIVAFSEFLR 478
Qy      469 LPY-----IHKAMVPRK--EKLKME 487
Db      479 GPYAKQTEPFPSKQPKKEIFIRKQKLE 506

```

RESULT 13

```

ID      08RMH4      PRELIMINARY;      PRT;      462 AA.
AC      08RMH4;
DT      01-JUN-2002 (TREMBLrel. 21, Created)
DT      01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Hypothetical protein.
GN      AT1G12640.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosida II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA      Palm C.J., Bowser L., Jones T., Barch J., Carinici P., Chen H.,
RA      Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA      Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sekano H.,
RA      Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
RA      Ecker J., Theologis A., Davis R.W.;
RL      Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA      Palm C.J., Bowser L., Jones T., Barch J., Carinici P., Chen H.,
RA      Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA      Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sekano H.,
RA      Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
RA      Ecker J., Theologis A., Davis R.W.;
RL      Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY093087; AAM13086.1; -
DR      EMBL; AY128762; AAM91162.1; -
DR      InterPro; IPR004299; MBOAT_fam.
DR      Pfam; PF03062; MBOAT; 1.
KW      Hypothetical protein.

```

SQ SEQUENCE 462 AA; 52151 MW; 3817A68E0CB35A5 CRC64;

Query Match 14.9%; Score 383.5; DB 10; Length 462;

Best Local Similarity 25.6%; Pred. No. 2.6e-24;

Matches 125; Conservative 86; Mismatches 215; Indels 59; Gaps 13;

```

Qy      26 LSLNLTATSLGSEQALRLIISIFLGYPFALFYRHLYPEKETYLIHLPHPTGSLAYFN 85
Db      1 MDMSMASIGSVAVLFLFCFVATIVPSACR---IVPSRLGHLVAAASGAVLSLIS 57
Qy      86 FGNQYHSLCLICVLOFLIRLMGRITITAV-----LTFCCQOMAYLAGYYTATGNY-- 137
Db      58 FG-----FSSNLHFLVPMITIGVASMAYIPKCGIITFFLGFALLICGHVFMGSDMK 110
Qy      138 --DIKMTPHCVLTLLKLGIAVDYFDGKDQNSLSEQOKAIRGVSLLEVAQSYFYG 195
Db      111 ECGIDSTALNVLTLKVLISCSKNYNDGMLKEGLEAKRNLQIOMPSLIRKFGCLCCG 170
Qy      196 AFLVGPQFSNMHYMKLVQGEI-----DIPGKIPIISIPALKRSLGLFYLAVGT 245
Db      171 SHFAPGVYEMQDYLEWTEGKGIWDTTERKKRPSYPGATIRAILQALICMALYLVLPQY 230
Qy      246 LISPHITEDYLTEDYDNHPFWRFCMYMLINGKFLYKYVTCMLTBEVCILTGLGNF 305
Db      231 L--TRFTPEPV-----YQEMGFPRKPSYQYMGFTARMKYFYIWSISEASIIISGDFSGW 283
Qy      306 EEKG--KAKMDACAMKVMLEPTNPRFTGSIASFNINNAVARYIFRKLKFNKE--- 360
Db      284 TDDASPKPRMDRANKVDILGVELAKSAVOIPLWMIQVSTWIRHYVYRELNVONGKAGFF 343
Qy      361 --LSQGLSLFLALMHGLSGYLVCFQMEFLIVEROAARLIQES--PTSKLAAITVL 416
Db      344 QLLATQVTS---AVMHGILPGYWMFVQSAIIMGSRVYIRMOQAIIPKAMLNINWVF 399
Qy      417 QPFYIVQOQTIHMLFMQYSMTAFCLFTWDMKLVKYSIYLGHIIFSL--FLIPIYH 473
Db      400 INFLYTV-----LVLNYSVAGFVWLSLHETLTAYGSVYIGTIIIPGLILSLYVVPAPK 453
Qy      474 KAMVPRKEK 482
Db      454 SRPKRKEE 462

```

RESULT 14

```

ID      018856      PRELIMINARY;      PRT;      480 AA.
AC      018856;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      C54G7.2 protein.
GN      C54G7.2.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nemtoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Pelodetinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=94150718; PubMed=7906398;
RA      Wilson R., Almscough R., Anderson K., Baynes C., Berka M.,
RA      Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA      Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA      Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA      Jones M., Kerhaw J., Kirsten J., Laister N., Latreille P.,
RA      Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,
RA      Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA      Smaiden N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA      Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA      Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans."
RT      Nature 368:32-38(1994).
RN      [2]

```

```

RP SEQUENCE FROM N.A.
RA Du 2.1;
RT "The sequence of C. elegans cosmid C54G7."
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40410; AAA81391.1; -.
DR WormBep; C54G7.2; CE04270.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
SQ SEQUENCE 480 AA; 55273 MW; 2229DA0ADDF5B55 CRC64;

Query Match 14.9%; Score 383.5; DB 5; Length 480;
Best Local Similarity 24.1%; Pred. No. 2.7e-24;
Matches 112; Conservative 98; Mismatches 221; Indels 33; Gaps 13;

QY 28 LNKLAATSLGSEQLRLIISIFLGPALFYRHYLFYKET--YLHLPHPTGSLIAYFN 85
DB 14 LEPLAETIGIADRVNVLISLFCFGISYVRKTCSLKQVNRQVTVPAVVGIVGIFFC 73
QY 86 FGNOLYHSLCTIVLQFLILRMGRT-ITAVLTTCFQMAVLLAGYYTTATGNDIKWTMP 144
DB 74 FGRALKILANALGSVAIMYFAPPTOVAKIVLLFSWGYLFEIHCRWALLESYSLDVTGP 133
QY 145 HCVLTLLTGLAVDPFG-GKDONSLSEOOKYAIRGVPISLLEVAGSFYGAFLVGPQF 203
DB 134 IMVAVEKATWMAFNKDGAKADSKLTBEQRESLKEIPSLSEMSFMENQIVLTGPAN 193
QY 204 SMNHMYKLVOGE-LIDIPKIPNSIIPALKRSLGLFYLL-VGYTLLSPHITBEDYLLTD 260
DB 194 NYVDYIFLDEKHLVADKHGLSPPTGAAMKKFYQSIFLAIIVTLGKKVVD-VGTGP 252
QY 261 YDNHPE--WRRCMYLMWKFVLYKYVTCMLVTEGVCITLGLNGEPEKGAADACAN 318
DB 253 YFALPFQWFFWFFITIF--FIRCAVYFAMWFADAINMSGFGSGYDKGNEMWKLCTN 310
QY 319 MKWALFPTNRFCTIISFNINTAMVARYIEKELKGLNGKELSGSLFLMLMHLGS 378
DB 311 VLPEVMAOSKELTLDGMNIQTGFMLKGYER---APKSIITVAATYLSAAMHGVSI 366
QY 379 GYLVCFQMEFLIVVERQARLQES-----PLTSKLAATVLAQFYVLAQOTIHLFM 432
DB 367 GYVWAFPTGFLFY---AAQTFRSMRWRLDHNKKFADI---FSFIISK---IAL 415
QY 433 GYSMTAFCLFTWDKMLKYKSIYFLGHIPLSLFLIPIYHKAM 476
DB 416 AYATYSFTWMMYPAVDVLKRVYFIPHIVAVAVIFALPKFEPKL 459

RESULT 15
Q9LN83 PRELIMINARY; PRT; 457 AA.
AC Q9LN83;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE T12C24.17;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Johnson-Hopson C., Khan S., Kim C.,
RA Shun P., Altafi H., Bei O., Chin C., Chio J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thavert A.,
RA Tortum M., Vayberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;

```

```

RT "Genomic sequence for Arabidopsis thaliana BAC T12C24 from chromosome
RT 1."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025417; AAF88094.1; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
SQ SEQUENCE 457 AA; 51600 MW; BB4E4BF2AA4BD CRC64;

Query Match 14.8%; Score 382.5; DB 10; Length 457;
Best Local Similarity 26.0%; Pred. No. 3.1e-24;
Matches 126; Conservative 82; Mismatches 217; Indels 59; Gaps 13;

QY 31 LATSLSGSEQLRLIISIFLGPALFYRHYLFYKET--YLHLPHPTGSLIAYFNGQL 90
DB 1 MASIISGVAVLFLICFVATIVSFACR---IVPSRLGHLVLAASGATLSLISFC--- 54
QY 91 YHSLCTIVLQFLILRMGRTIV---LTPFCQMAVLLAGYYTTATGNY---DIK 140
DB 55 ---FSSNLHFLVPMITGYASMAIYRPKCGIITFFLGFAYLICHVFMGDMKEGID 110
QY 141 WTMPHCVLTLLGLAVDPFGGKDONSLSEOOKYAIRGVPISLLEVAGSFYGAFLV 200
DB 111 STALMVLTLKVLISCSKMYNDGMLKEGLREAOQRRLIOMPSLIEYFGCLCCGSHFAG 170
QY 201 POFSMNHMYKLVOGELI-----DIPKIPNSIIPALKRSLGLFYLVGYTLLSPH 250
DB 171 PYEMDYLEMTGSKGIWTEKRKKRSPYGATIRALQALICMALVLYVQYPL--TR 228
QY 251 ITDLYLTEDYDNHPEWRCMYLMWKFVLYKYVTCMLVTEGVCITLGLNGEPEKGA 309
DB 229 FTEPV---YQWGFARKEPSYQYMGFTARMKYYFPIWSISEASIIISGLGFSGMTDAS 283
QY 310 -KKKWDACAMKWLPTNRFCTIISFNINTAMVARYIEKELKGLNGKELSGSLFLML 362
DB 284 PKPMDAKKVDILGVELAKSAVOIPLWNIQVSTLRHAYELVONGKAGAFQOLLAT 343
QY 363 QGSLFLALMHLGSLGYLVCFQMEFLIVVERQARLQIE-SEPLSKLAATVLAQFY 421
DB 344 QTVS---AVMHGILPYWMPFYQSALMINGSVITYWQQAISPKMALNINWIFINFLY 399
QY 422 LVQOTIHLFMGYSMTAFCLFTWDKMLKYKSIYFLGHIPLSL---FLIPIYHKAMP 478
DB 400 TV-----LVNYSAGVFMVLSHETLTAYGSYYIGTIIPVGLILSYVAPAKRBRKP 453
QY 479 RKEX 482
DB 454 RKEB 457

```

Search completed: November 21, 2003, 14:16:04
 Job time : 41 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:02:07 ; Search time 47 Seconds
(without alignments)
1644.677 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 2580
Sequence: 1 MASSAEGDEFTVALAGVLQ.....ILPYIKAMPKREKTKME 487

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2580	100.0	487	21	AA169987
2	2580	100.0	487	22	AA169987
3	2580	100.0	487	22	AA169987
4	2265	87.8	424	22	ABG06198
5	871.5	33.8	497	22	ABB71792
6	638	24.7	128	21	AA169987
7	408.5	15.8	722	22	AB169987
8	407	15.8	465	23	AAU72527
9	389	15.1	520	24	ABU19323

10	362.5	14.1	487	24	ABU19324
11	334.5	13.0	432	22	AA169987
12	334.5	13.0	432	22	AA169987
13	328.5	12.7	423	22	AA169987
14	241.5	9.4	489	22	AB169987
15	231	9.0	256	23	AB169987
16	215	8.3	58	22	AB169987
17	215	8.3	58	22	AB169987
18	215	8.3	58	22	AB169987
19	215	8.3	58	22	AB169987
20	215	8.3	58	22	AB169987
21	215	8.3	58	22	AB169987
22	215	8.3	58	22	AB169987
23	215	8.3	58	22	AB169987
24	215	8.3	58	22	AB169987
25	215	8.3	58	22	AB169987
26	184.5	7.2	345	22	AB169987
27	177	6.9	371	22	AB169987
28	174.5	6.8	467	22	AB169987
29	173	6.7	799	22	AB169987
30	170	6.6	429	22	AB169987
31	166.5	6.5	275	22	AB169987
32	166.5	6.5	424	22	AB169987
33	163	6.3	139	23	AB169987
34	156.5	6.1	141	22	AB169987
35	148	5.7	267	22	AB169987
36	142	5.5	525	22	AB169987
37	133	5.2	453	23	AB169987
38	129.5	5.0	418	23	AB169987
39	124.5	4.8	112	23	AB169987
40	122	4.7	493	22	AB169987
41	122	4.7	493	22	AB169987
42	115	4.5	222	23	AB169987
43	114	4.4	153	23	AB169987
44	113	4.4	137	21	AB169987
45	112.5	4.4	404	21	AB169987

ALIGNMENTS

RESULT 1
AA169987
ID AA169987 standard; Protein, 487 AA.

AC AA169987;
XX 31-MAY-2000 (first entry)

Human receptor-associated protein from Incyte clone 2906971.

Human receptor-associated protein; HRAP; Incyte clone 2906971;
cytostatic; immunomodulatory; antiinflammatory; cardiant; antiaemic;
antiarteriosclerotic; hepatotropic; antiatherosclerotic;
antiaesthetic; osteopathic; antiallergic; antidiabetic; dermatological;
neuroproliferative; diagnosis; treatment; prevention; reproductive disorder;
cardiovascular; cell proliferative; autoimmune; inflammatory; allergy;
gastrointestinal; atherosclerosis; cirrhosis; leukaemia; cancer; AIDS;
arthritis; anaemia; asthma; dermatitis; diabetes; osteoporosis;
multiple sclerosis; irritable bowel syndrome.

XX Homo sapiens.

OS Homo sapiens.

Key Location/Qualifiers
Modified-site 3
/note= "Potential phosphorylation site"
Modified-site 183
/note= "Potential phosphorylation site"
Modified-site 149
/note= "Potential phosphorylation site"
Modified-site 235
/note= "Potential phosphorylation site"

NOVX related prote
Human stomach can
Human protein sequ
Human AFP protein
Drosophila melanog
Human secretory po
Human liver peptid
Human peptide #119
Peptide #116 endo
Human brain expres
Human bone marrow
Peptide #122 endo
Peptide #120 endo
Human peptide enco
Human hydrophobic
Novel human diagno
Human polypeptide
Novel human diagno
Human polypeptide
Human EST encoded
Novel human diagno
Human ORFX protein
Novel human diagno
Novel human diagno
Drosophila melanog
Human malignant me
Streptococcus poly
Arabidopsis parlia
Human protein sequ
Novel human protei
Human ovarian anti
Human ORX protein
S. xyloisus D1B pr

PN WO200008155-A2.
 XX 17-FEB-2000.
 XX
 XX 06-AUG-1999; 99WO-US17777.
 XX
 XX 07-AUG-1998; 98US-0160065.
 PR 01-SEP-1998; 98US-0098703.
 XX
 XX (INCY-) INCYTE PHARM INC.
 XX
 PI Hillman JL, Yue H, Lai P, Tang YT, Gorgone GA, Guegler KJ;
 PI Corley NC, Baughn MR;
 DR WPI: 2000-205710/18.
 DR N-PSDB; AAF250889.
 XX
 XX New human receptor-associated proteins (HRAP) useful for the diagnosis,
 PT treatment and prevention of cell proliferative, autoimmune,
 PT inflammatory, reproductive, cardiovascular, and gastrointestinal
 PT disorders -
 PS Claim 1; Pages 73-74; 99pp; English.
 XX
 XX The present sequence is a human receptor-associated protein
 CC (HRAP) from Incyte clone 2906971 obtained from THYMOT05 cDNA library.
 CC This sequence is expressed in nervous, gastrointestinal
 CC and reproductive tissues. HRAP has cytostatic, immunomodulatory,
 CC antiinflammatory, cardiac, antiarteriosclerotic, hepatotrophic,
 CC antitumorigenic, antineumatic, osteopathic, antiallergic, antianemic,
 CC antiasthmatic, antidiabetic, dermatological and neuroprotective
 CC activities. The present sequence is useful in the diagnosis, treatment
 CC and prevention of disorders associated with HRAP expression, especially
 CC cell proliferative, autoimmune/inflammatory, reproductive,
 CC cardiovascular and gastrointestinal disorders (e.g. atherosclerosis,
 CC cirrhosis, leukaemia, cancer, AIDS, arthritis, allergies, anaemia,
 CC asthma, dermatitis, diabetes, osteoporosis, multiple sclerosis and
 CC irritable bowel syndrome).
 XX
 XX Sequence 487 AA;
 SO
 Query Match 100.0%; Score 2580; DB 21; Length 487;
 Best Local Similarity 100.0%; Pred. No. 5.5e-275;
 Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 YLVQOTIHLFWGYSMTAFCLFTWDMKLVKYSIYFLGHIFLSTFLPYIHKAMVRK 480
 DB 421 YLVQOTIHLFWGYSMTAFCLFTWDMKLVKYSIYFLGHIFLSTFLPYIHKAMVRK 480
 QY 481 EKLKQME 487
 DB 481 EKLKQME 487
 RESULT 2
 AAB3169
 ID AAB3169 standard; Protein; 487 AA.
 XX
 AC AAB3169;
 XX
 DT 30-APR-2001 (first entry)
 XX
 XX Amino acid sequence of a human protein having a hydrophobic domain.
 DE Human; hydrophobic protein; secretory protein; membrane protein; sepsis;
 XX tumour inhibition; immune deficiency; autoimmune disorder; anaemia; burn;
 KW infectious disease; cancer; ulcer; periodontal disease; coagulation;
 KW Parkinson's disease; fertility; immune response; thrombosis.
 XX
 XX Homo sapiens.
 OS
 PN WO200104297-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 16-JUN-2000; 2000WO-JP03942.
 XX
 PR 08-JUL-1999; 99JP-0194359.
 XX
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;
 XX
 PS WPI: 2001-103081/11.
 DR N-PSDB; AAF25159; AAF25169.
 XX
 PT Isolated human proteins and polynucleotides are used in research and
 PT have activities including cell proliferation/differentiation activity,
 PT immune stimulating activity and receptor/ligand activity -
 XX
 XX Claim 1; Page 90-93; 151pp; English.
 PS
 XX The present sequence represents a human protein with hydrophobic domains.
 CC The protein possesses a hydrophobic domain and so is a secretory protein
 CC or a membrane protein. The protein is used as an antigen to prepare
 CC antibodies. The polynucleotide sequence is useful as a source of probes
 CC for genetic diagnosis. It is also useful for producing the protein
 CC in large quantities and for gene therapy. The eukaryotic cells are used
 CC for detecting the receptors or ligands corresponding to the protein and
 CC for detecting small novel pharmaceuticals. The antibodies are also used
 CC for detection, quantification and purification of the proteins. Both the
 CC protein and polynucleotide may be used in research or as nutritional
 CC sources or supplements. The protein may have cytokine and cell
 CC proliferation/differentiation activity, immune stimulating or suppressing
 CC activity, hematopoiesis regulating activity, tissue growth activity,
 CC and thrombolytic activity, chemotactic/chemokinetic activity, hemostatic
 CC activity and tumour inhibition activity. It may therefore be used to
 CC treat immune deficiencies resulting from autoimmune disorders or
 CC infectious diseases, cancer, sepsis, anaemias, burns and ulcers,
 CC periodontal disease, Parkinson's disease, induce fertility, improve
 CC immune response and enhance coagulation or inhibit thrombosis.
 XX
 SO Sequence 487 AA;
 Query Match 100.0%; Score 2580; DB 22; Length 487;
 Best Local Similarity 100.0%; Pred. No. 5.5e-275;

QY 481 EKLKME 487
 |||||
 Db 481 EKLKME 487

RESULT 4
 ABG06198
 ID ABG06198 standard; Protein; 424 AA.
 XX

AC ABG06198;

XX 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6189.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM Food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS70385.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20; SEQ ID NO 36557; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical.
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostic, forensic, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 424 AA;

Query Match 87.8%; Score 2265; DB 22; Length 424;
 Best Local Similarity 99.5%; Pred. No. 2.3e-240;
 Matches 422; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 64 YKEHYLIHLEHFTGTGLSIAYFNGNQLVHSLLCIVLQFLILRLMGRITAVLTTCFQMA 123
 |||||

Db 1 YKEHYLIHLEHFTGTGLSIAYFNGNQLVHSLLCIVLQFLILRLMGRITAVLTTCFQMA 60
 QY 124 YLLAGVYVTAANGNDIKMTMPHCVLTLKLGIAVYFDGKQNSLSEQKXAIKGVPS 183
 |||||
 Db 61 YLLAGVYVTAANGNDIKMTMPHCVLTLKLGIAVYFDGKQNSLSEQKXAIKGVPS 120
 QY 184 LLEVAGFSYFYGAFLVGPQFSNMHYMKLVQGLIDIPKIPNSIIPALKRSLSLGFYLVG 243
 |||||
 Db 121 LLEVAGFSYFYGAFLVGPQFSNMHYMKLVQGLIDIPKIPNSIIPALKRSLSLGFYLVG 180
 QY 244 YTLSPHITEDYLLTEDYDNDHPEWRCWMLIWGFVLKYKVTCLVTEGVCITLGLGN 303
 |||||
 Db 181 YTLSPHITEDYLLTEDYDNDHPEWRCWMLIWGFVLKYKVTCLVTEGVCITLGLGN 240
 QY 304 GFEKGRKAKWDCAMMKWLPETNPRFGTASFININAMARITPRLKFLNGKELSQ 363
 |||||
 Db 241 GFEKGRKAKWDCAMMKWLPETNPRFGTASFININAMARITPRLKFLNGKELSQ 300
 QY 364 GLSLFLALMHGLHSGYLVCFQMEFLIYVERQARLLIQESFTLSKLAITVLQPFYLV 423
 |||||
 Db 301 GLSLFLALMHGLHSGYLVCFQMEFLIYVERQARLLIQESFTLSKLAITVLQPFYLV 360
 QY 424 QQTIMWLFMGYSMTAFCLFTWDKWLKVKSIYFLGHIFFLSLFLPYIHKAMVPRKEXL 483
 |||||
 Db 361 QQTIMWLFMGYSMTAFCLFTWDKWLKVKSIYFLGHIFFLSLFLPYIHKAMVPRKEXL 420
 QY 484 KME 487
 |||||
 Db 421 KME 424

RESULT 5
 ABB71792
 ID ABB71792 standard; Protein; 497 AA.
 XX
 AC ABB71792;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 42168.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-65660/75.
 DR N-PSDB; ABL15895.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Disclosure; SEQ ID NO 42168; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

Sequence 497 AA:

Query Match 33.8%; Score 871.5; DB 22; Length 497;
 Best Local Similarity 37.2%; Pred. No. 11e-86;
 Matches 167; Conservative 92; Mismatches 165; Indels 25; Gaps 5;

QY 28 LNKATSLGASEQALRLIISIFLGFALFYRHV-FKETYLIHLFTFTGLSIAYNF 86
 Db 14 MDGIASGVPEALRLITLTLGYPVALVQKFSIADKVNHFAGGAGLCYFNY 73
 QY 87 GNOYLHSLCTV-LQFLRLMGRITVLTTCFQMAVYLAGYVYTAQNDIKWMPH 145
 Db 74 GLDYHSLIALITLYFVLRLRKTKQIFLAINFVFMGYLLIGFYTSNDYDILMTMPH 133
 QY 146 CVLTLKLGLAVDYFDGKDNLSSEQKVAIRGVPSLLFVAGSYFYGAFLVGPQFSM 205
 Db 134 CILVLRMIGYGFDTDGLKESELSKDKETALKRPSLLBELLAFSYPPSGFLVGPQFP 193
 QY 206 NHMYKLVGELIDIPGKIIPNSIIPALKRLSLGLFYLVGYTLLSPHITEDYLLTEDYDNR 265
 Db 194 RRYAFVDFGRFHNGENVAGV---RRFGAGAFILYICQGLRKLPSYFLTEPFAQVS 249
 QY 266 FWFRCMYLWNGKFLVLYRYTCWLVTEGVCILTGLGNGFEBKAKAKADACANMKWLF 325
 Db 250 FVKRIYLLGFPAKFSLYKYISCMILTEGALICIGLYKGEIDKNGQPDWGSQSNVTKLLE 309
 QY 326 TNPRTGTIASFNINTNMAVARYLFRKLKFLGNKELSGSLPLALMHGHSGLVCFQ 385
 Db 310 TGNTEHMYVQSNVNTQWVGQYIKRLKFLNNRTISGALGFVAVHGHSGYMTFL 369
 QY 386 MEFLIVERQAARL-----IOESPLSLAATVLOPFYVVOQTIHMLFMGY 434
 Db 370 MEYMWVSTEKQITRYTQVLPQMGHILNNSDIYTL-----LYFTLLKSYNVVYNGW 421
 QY 435 SMTAFCLFTWDKMLKVYSIYFLGHIIFL 463
 Db 422 CLTAFFVFLKRYERMIYVGAVSYGFTFLV 450

RESULT 6
 AAG00235
 ID AAG00235 standard; Protein; 128 AA.

XX AAG00235;
 XX
 XX AAG00235;
 XX
 XX 06-OCT-2000 (first entry)
 XX
 XX Human secreted protein, SEQ ID NO: 4316.
 XX
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 XX Homo sapiens.
 OS
 XX
 XX EP1033401-A2.
 PN
 XX
 XX 06-SEP-2000.
 PD
 XX
 XX 21-FEB-2000; 2000EP-0200610.
 PF
 XX
 XX 26-FEB-1999; 99US-0122487.
 PR
 XX
 XX (GBST) GENSET.
 PA
 XX
 XX Dumas Milne Edwards J, Duclet A, Giordano J;
 FI
 XX

DR WPI; 2000-500381/45.
 DR N-PSDB; AAC00241.
 XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 XX
 XX Claim 13; SEQ ID 4316; 71pp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 CC
 XX
 XX Sequence 128 AA;

Query Match 24.7%; Score 638; DB 21; Length 128;
 Best Local Similarity 99.2%; Pred. No. 7.7e-62;
 Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MASSAEDDEGVVALAGVLOGSPQELSLNKATSLGASEQALRLIISIFLGFALFYRH 60
 Db 1 MASSAEDDEGVVALAGVLOGSPQELSLNKATSLGASEQALRLIISIFLGFALFYRH 60
 QY 61 YLFKETYLIHLFTFTGLSIAYNFGNOLYHSLCTVLOFLRLMGRITTAVALTTCF 120
 Db 61 YLFKETYLIHLFTFTGLSIAYNFGNOLYHSLCTVLOFLRLMGRITTAVALTTCF 120
 QY 121 QMAVLAG 128
 Db 121 QMAVLAG 128

RESULT 7
 ABB66994
 ID ABB66994 standard; Protein; 722 AA.

XX ABB66994;
 XX
 XX 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 27774.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX
 XX WO200171042-A2.
 PN
 XX
 XX 27-SEP-2001.
 PD
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR
 XX
 XX 11-JUL-2000; 2000US-0614150.
 PR
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX
 XX WPI; 2001-656860/75.
 DR
 XX
 XX N-PSDB; ABL11097.
 DR
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 interactions -

PS Disclosure; SEQ ID NO 27774; 21np + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB872072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 722 AA:

Query Match 15.8%; Score 408.5; DB 22; Length 722;

Best Local Similarity 25.8%; Pred. No. 2e-35; Mismatches 212; Indels 47; Gaps 14;

Matches 121; Conservative 89; Mismatches 212; Indels 47; Gaps 14;

QY 31 IATSGASBOALRLIISIFLGPALFYRHYLFYKE--TYLIHLFHTFTGLSIAYFNFGN 88

DB 22 LADWGLSDVADNPLICQISALFLASLFRSMHPSVSKLHRTFALSGLAFGFCFQ 81

QY 89 QLYH----SLICIVQFLILRLMGRTTAVLTTF--CFQAYVLLAGYYTATGNDIKW 141

DB 82 QAIHTAGLPAICYIVRTDPRIVQRAVLVMSYLLCVHLMROLVDY----GSYALDI 136

QY 142 TMRPHCVTLKLGLAVDFDG--GKDONSLSSEQKAIKRGVSLLEVAGEFYAGFLNG 200

DB 137 TGRMLITTKVSLASFISHDGFEVGRDEELTKAQOYHAIKMPASLEYSVVHFGSILAG 196

QY 201 POFNMNHYKLVQ--ELIDIP--GKIPNS-----IIP--ALKRLSIG--LFLVVG 243

DB 197 PLVFKYDIEFVGVNLTSTPPGNGLDSSKREVLPEPTKAVIRKVGSLVCAFIEMK 256

QY 244 YTLSP--HITDYLITEDYDHNHPFRCMYLNGKEVLYKYVTCWLTBEGVCLITGIG 301

DB 257 FVKIYFVKCMKEDDFNMNTSMYKYV---YAMMATTCIRFYHAMLADAINNSGIG 312

QY 302 FNGFEKSKAKKADACANMKVWLPETPRFTGTIASNINTMAMVARYIKRLKFLGNKL 361

DB 313 FTGYDGDGSKMDLISNINVLSPSTNRDAIMNNNCGTNRWLTLYVERVP---QQY 368

QY 362 SOGLSLFLALMHGLSHGYLVCFOEFLIVIERQARLIQESPTLSKLAITVLQPFY 421

DB 369 GTLLTFALSAVWHGFPYGYLFPATGAVVTAARTGRILFRH----RQSTQVTRMFYD 423

QY 422 LVQQTTHMLFMGYSMATFCLFTWMDKLYKYKSIYFLGHIFFLSLFIPL 470

DB 424 ILLCLTRVVLGYATFPFVILFEMSGIKLYLRFYLCIHTISLVTIFILP 472

RESULT 8

AAU72527

ID AAU72527 standard; Protein; 465 AA.

AAU72527;

26-FEB-2002 (first entry)

Arabidopsis cell cycle protein CCP14.

Cell cycle protein; CCP; cell cycle regulation; herbicide;

plant growth regulator; plant development; abiotic stress; biotic stress;

nutrient deprivation; pathogen attack; crop yield; immunogen.

Arabidopsis thaliana.

PM WO200185946-A2.

XX 15-NOV-2001.

XX 14-MAY-2001; 2001WO-IB01307.

XX 12-MAY-2000; 2000US-204045P.

XX (CROP-) CROPPEDSIGN NV.

XX Inze D, Boudolf V, De Veylder L, Acoستا JAT, Magyar Z;

XX WPI; 2002-062249/08.

XX N-PEDB; AAS96317.

PT New cell cycle protein and nucleic acid molecule encoding it useful for

PT regulating cell cycle progression in plants and for identifying

PT modulators which are useful as herbicides or plant growth regulators -

PS Claim 34; Fig 14; 316pp; English.

CC The invention relates to a novel cell cycle protein (CCP) and the

CC polynucleotides encoding them. CCP is useful for identifying a compound

CC which modulates the activity of the polypeptide and which binds to the

CC polypeptide and an anti-CCP antibody is useful for detecting the presence

CC of CCP in a sample. A CCP modulator is useful for modulating the cell

CC cycle or growth of a plant such as *Arabidopsis thaliana*, rice, wheat,

CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.

CC CCP nucleic acid and polypeptide molecules are useful as modulating

CC agents in regulating cell cycle progression in plants. CCP is useful to

CC treat disorders characterized by insufficient or excessive production of

CC CCP protein or production of CCP protein forms which have decreased or

CC aberrant activity. Compounds that bind to or modulate the activity

CC of CCP polypeptide are useful as herbicides or plant growth regulators.

CC The polynucleotide is useful for modifying cell fate, plant development,

CC plant morphology, biochemistry and/or physiology, the length of the G1,

CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,

CC stimulation or enhancement of cell division, DNA replication, seed set,

CC seed size, seed development, tuber, fruit, leaf formation, shoot and root

CC initiation and/or development, nodule function, dwarfism in plants,

CC senescence, tolerance or resistance to stresses. CCP, the polynucleotide

CC and the anti-CCP antibody are useful in agriculture to modulate the

CC protein levels or activity of a protein involved in the cell cycle due

CC to environmental conditions, including abiotic stress such as

CC cold, nutrient deprivation, heat, drought, salt stress, or biotic

CC stress such as pathogen attack, to modulate e.g. enhance crop yields,

CC and attenuate plant architecture, plant quality traits, plant

Sequence 465 AA:

Query Match 15.8%; Score 407; DB 23; Length 465;

Best Local Similarity 25.4%; Pred. No. 1.5e-35;

Matches 123; Conservative 90; Mismatches 227; Indels 44; Gaps 12;

QY 26 ISLNLATSLGASBOALRLIISIFLGPALFYRHYLFYKE--TYLIHLFHTFTGLSIAYFN 85

DB 4 LDMNSMAISIGSVAVNLFPLCPVATITISPLMR---FIPSRGLGHITSAASGAPLSLUS 60

QY 86 FGNOLYHSLICIVQFLILRLMGRTTAV-----LTFCEQMAVLLAGYYTATGNY-- 137

DB 61 FG-----FSSNHFHFLPMTIGVSMAYPLSGSFTIFPLGFAVLICGHVFMVSGDAWK 113

QY 138 --DIKTMPHCVTLKLGLAVDFDGKDNSSLSSEQKAIKRGVSLLEVAGSYRYG 195

DB 114 EGGIDSTGALMVLTLKVISINVDGMLKEEGLEAGKQKRLIOMPSLIBYFGCLCG 173

CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 432 AA;

Query Match 13.0%; Score 334.5; DB 22; Length 432;
 Best Local Similarity 25.7%; Pred. No. 1.3e-27;
 Matches 98; Conservative 72; Mismatches 163; Indels 49; Gaps 12;

QY 118 FCFQNAVY-----LAGYYTATGNYDIKWTMPHCVLTLKLGAVDPDG--GKQNSLSS 171
 DB 15 FVFALGYLTVCQVTRVYIFDYGYSDAFSGPMWITQKITSLACEIHDMFRKDE-ELTS 73
 QY 172 EQOKAIVGPELLEVAGSFYFGAFLVGPFSNMHMYKLVOGELIDP-----GK----- 222
 DB 74 SQRDLAVRMPSLLEVLISYCNFMGILAPLCSYXDYIFFGRSYHITQSGENGKEETQ 133
 QY 223 -----IPNSIIPALKRLSLGLFYLVGYTLSPHITEDYLTLEDYDNHPFM-FRCMYMLI 275
 DB 134 YRTEPSPRTAV-VQKLIVCGLSLFLHITCTT-LPVEYNIDENHFOATSWPTKIIYLYI 191
 QY 276 WGFVLYKYVTCWLVTEGVCILTGLFGNGFEKGAKWDACANMKWLFETNPRFTGTIA 335
 DB 192 SLAARPKYFFAMTLADINNAAGFGFRGYDENGARWDLISNLRIQIEMSTSFEMFLD 251
 QY 336 SFNINTNMAVAYIFKRLKFLGNKELSGSLFLALMHGHSGLVCGFQMEFLIYVER 395
 DB 252 MNNIQTALMLKVKVCERTSFSPITQ-----TPILSAIMHGVYPGYLFLTGVLMTL--- 303
 QY 396 QARLIQES-----PTLSKLAATVLOPFYLVQOITIMLPMGYSMTAFCLFTWDKWL 448
 DB 304 -AARARNRNFHYFLFEPQKL-----FYDVTIWTVOVAISTVVPVLLSIKPSL 354
 QY 449 KYVKSIFYLGHIFPLSLFLIP 470
 DB 355 TFYSSWYICHLILGLVLLLP 376

RESULT 13

AAAG81345
 ID AAG81345 standard; Protein; 423 AA.

AC AAG81345;

DT 10-SEP-2001 (first entry)

DE Human AFP protein sequence SEQ ID NO:208.

XX Human; secreted protein; secretion; bacterial cell; fungal cell;
 XX eukaryotic cell; fusion protein; maltose binding protein;
 KW immunoglobulin constant region; polynucleotide tag.

OS Homo sapiens.

PN WO200129221-A2.

PD 26-APR-2001.

PF 20-OCT-2000; 2000WO-US29052.

PR 20-OCT-1999; 99US-0160712.

PA (ZYMO) ZYMOGENETICS INC.

XX

PI Conklin DC, Yee DP;
 XX WPI; 2001-300340/31.
 DR N-PSDB; AAH52196.

PT Isolated polypeptide for directing secretion of proteins of interest
 PT from a host cell including, e.g. bacteria, includes contiguous amino
 XX acid residues of polypeptide with specified amino acids
 PS Claim 1; Page 362-364; 617pp; English.

XX AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242
 CC to AAG81453. The secreted proteins can be used for directing the
 CC secretion of proteins of interest from a host cell including bacteria,
 CC fungal cells, and cultured higher eukaryotic cells. The present invention
 CC also describes fusion proteins, where a secreted protein of the invention
 CC is operably linked via a peptide bond or peptide linker to a second
 CC protein selected from the group consisting of maltose binding protein,
 CC an immunoglobulin constant region, a polynucleotide tag and a peptide
 CC given in AAG81453.

XX Sequence 423 AA;

Query Match 12.7%; Score 328.5; DB 22; Length 423;
 Best Local Similarity 25.4%; Pred. No. 5.9e-27;
 Matches 97; Conservative 72; Mismatches 164; Indels 49; Gaps 12;

QY 118 FCFQNAVY-----LAGYYTATGNYDIKWTMPHCVLTLKLGAVDPDG--GKQNSLSS 171
 DB 6 FVFALGYLTVCQVTRVYIFDYGYSDAFSGPMWITQKITSLACEIHDMFRKDE-ELTS 64
 QY 172 EQOKAIVGPELLEVAGSFYFGAFLVGPFSNMHMYKLVOGELIDP-----GK----- 222
 DB 65 SQRDLAVRMPSLLEVLISYCNFMGILAPLCSYXDYIFFGRSYHITQSGENGKEETQ 124
 QY 223 -----IPNSIIPALKRLSLGLFYLVGYTLSPHITEDYLTLEDYDNHPFM-FRCMYMLI 275
 DB 125 YRTEPSPRTAV-VQKLIVCGLSLFLHITCTT-LPVEYNIDENHFOATSWPTKIIYLYI 182
 QY 276 WGFVLYKYVTCWLVTEGVCILTGLFGNGFEKGAKWDACANMKWLFETNPRFTGTIA 335
 DB 183 SLAARPKYFFAMTLADINNAAGFGFRGYDENGARWDLISNLRIQIEMSTSFEMFLD 242
 QY 336 SFNINTNMAVAYIFKRLKFLGNKELSGSLFLALMHGHSGLVCGFQMEFLIYVER 395
 DB 243 MNNIQTALMLKVKVCERTSFSPITQ-----TPILSAIMHGVYPGYLFLTGVLMTL--- 294
 QY 396 QARLIQES-----PTLSKLAATVLOPFYLVQOITIMLPMGYSMTAFCLFTWDKWL 448
 DB 295 -AARARNRNFHYFLFEPQKL-----FYDVTIWTVOVAISTVVPVLLSIKPSL 345
 QY 449 KYVKSIFYLGHIFPLSLFLIP 470
 DB 346 TFYSSWYICHLILGLVLLLP 367

RESULT 14

ABB64492
 ID ABB64492 standard; Protein; 489 AA.

AC ABB64492;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 20268.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX

PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PMD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL08595.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 20268; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (AB01840-ABL16175) and the encoded proteins
CC (AB057737-AB072072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 489 AA;

Query Match 9.4%; Score 241.5; DB 22; Length 489;
Best Local Similarity 23.1%; Pred. No. 2,8e-17;
Matches 107; Conservative 73; Mismatches 195; Indels 89; Gaps 17;

QY 77 TGLSIAYFNFGNQLVHSLICIVLQFLILMLMGRITTAV-----LTFECFQMAVL-----L 126
DB 36 TGLSLVAVVIVVYSGL-HSLHC-----FVSLALGTAVALVHPSGHLVTFVMMGVLFVRFRI 90
QY 127 AGYYTATGANGYDIKWTMPHCIVTLKLIQLAVDFDQGDQNSLSSBOQKTAIRGVP----- 182
DB 91 FDFYFGIRGHTN-----MIQMILTLKVSGLAFKETAAMK-RLQAHDEQKKNQDRDVGQESP 145
QY 183 -----SLEVAGSEYFYGAFVVGQFSKNNHYMKLVQGLIDIPKIPNSITP 229
DB 146 IEITDYDELQSLSAEILHVSFNYIGVLTGPYRRTYR-----DYFEMPRTYAPFVE 200
QY 230 A-LKRLSLGLFLVGYTLLSPHITEDYLLTECYDN-HPEWRCQMLIMGKFLVLYKYVC 287
DB 201 ALELKLKAVFYCALYLATNYMPDYLALSDFFDRSDVRYLLY--VWPTFTFRARIY 258
QY 288 WLVTGVCILTLGLGFGNGFEKCG-----KAKMDA-----CANMKVWLFE 325
DB 259 TGLTISECCCTMAAGGAYPDESDPNNGBGRKRYOHLKRDADKHNYNTFTTYNTVLEVE 318
QY 336 TNPRTGTITASNINNTANVARIYIFKRLKELCNKELISGLSLFLALMHLGSHVLCFQ 385
DB 319 RCMTFREGKHMNVCCQYWLAVNVY---KLFPSSKYRTGATLLCAAYWHGFRPHYFCIM 375
QY 366 MEFLVIYERQAAARLLOESPTLSKLAITVLOPFYLVQOTTHMEFMGSMN-----APCL 441
DB 376 GAPFYVSLDDMDKLVKRSATGTSRRVIV-----LFWIFKMFATFVLSGAEFL 424
QY 442 FTWDMKLVKYSIYFGLHIFFLSL---FILPYIHKAWPRKEX 482
DB 425 SSFGIMWRYSVYHIGYISMAAMTALGTYLSQKRAARRRKR 468

RESULT 15
ABB97850

ID ABB97850 standard; Protein; 256 AA.
XX
XX ABB97850;
AC
XX
XX 03-OCT-2002 (first entry)
DT
XX
XX Human secretory polypeptide (SPTM) 102.
DE
XX
XX Human; secretory protein; secretory polynucleotides; SPTM;
KW severe combined immunodeficiency; intracellular parasite protection;
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
KW motor neuron disorder; demyelinating disease; multiple sclerosis;
KW meningitis; abscess; prion diseases; cerebral palsy;
KW neuroskeletal disorder; peripheral nervous system disorder;
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
KW mental disorder; Tourette's syndrome.
XX
XX Homo sapiens.
OS
XX
XX WO200220756-A2.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 30-AUG-2001; 2001WO-US27297.
PF
XX
XX 05-SEP-2000; 2000US-229747P.
PR 05-SEP-2000; 2000US-229748P.
PR 05-SEP-2000; 2000US-229749P.
PR 05-SEP-2000; 2000US-229750P.
PR 05-SEP-2000; 2000US-229751P.
PR 05-SEP-2000; 2000US-230016P.
PR 05-SEP-2000; 2000US-230018P.
PR 06-SEP-2000; 2000US-230505P.
PR 06-SEP-2000; 2000US-230514P.
PR 06-SEP-2000; 2000US-230515P.
PR 06-SEP-2000; 2000US-230517P.
PR 06-SEP-2000; 2000US-230518P.
PR 06-SEP-2000; 2000US-230519P.
PR 06-SEP-2000; 2000US-230595P.
PR 06-SEP-2000; 2000US-230596P.
PR 06-SEP-2000; 2000US-230597P.
PR 06-SEP-2000; 2000US-230599P.
PR 06-SEP-2000; 2000US-230610P.
PR 06-SEP-2000; 2000US-230864P.
PR 06-SEP-2000; 2000US-230865P.
PR 06-SEP-2000; 2000US-230988P.
PR 06-SEP-2000; 2000US-230989P.
PR 06-SEP-2000; 2000US-230990P.
PR 07-SEP-2000; 2000US-230897P.
PR 07-SEP-2000; 2000US-230951P.
PR 07-SEP-2000; 2000US-231163P.
PR 07-SEP-2000; 2000US-231832P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL,
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR,
PI Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A,
PI Marabha R, Chen AJ, Chang SC, Au AP, Imman RR;
XX WPI; 2002-315658/35.
DR N-PSDB; ABL99847.
XX
XX Polynucleotide sequences encoding human secretory proteins useful for
PT gene therapy of e.g. genetic deficiency disorders, cancers, and
PT diseases caused by intracellular parasites -
XX
XX Claim 29; Page 430-431; 585pp; English.

CC The invention comprises the amino acid and coding sequences of human
 CC secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are
 CC useful for treating a disease or condition associated with the expression
 CC of functional SPTM. The SPTM DNA sequences are useful for somatic or
 CC germline gene therapy to correct a genetic deficiency (e.g. severe
 CC combined immunodeficiency). The SPTM DNA sequences are also useful in
 CC providing protection against intracellular parasites (e.g. fungal
 CC parasites and protozoan parasites). The SPTM DNA and protein sequences
 CC are also useful for diagnosing cell proliferative disorders, cancer,
 CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
 CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
 CC sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
 CC neuroskeletal disorders, peripheral nervous system disorders,
 CC dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
 CC disorders (e.g. Tourette's syndrome). Amino acid sequences ABB97749 -
 CC ABB97933 represent human secretory proteins of the invention.

XX Sequence 256 AA;

Query Match 9.0%; Score 231; DB 23; Length 256;
 Best Local Similarity 26.3%; Pred. No. 1.5e-16;
 Matches 65; Conservative 43; Mismatches 125; Indels 14; Gaps 4;

QY 234 LSLGLFYLVGYTLSPHITEDYLLTDYDNNHPFRCMYMLWGKFLYKYTCMLVTEG 293
 DB 11 LSLLEFLTLTKTFPPVTCVVDWDFV---HKASFPARLCYLVVMQASKPKYFAWTLADA 66
 QY 294 VCLITGLGNGFEEKAKADACANMKVLFETNPRFTGTIASFNINTANAVARIYFKRL 353
 DB 67 VNNAGFGSGVDKNGNFCWDLNLINIKIETATSKMYLENMNIQTATWLCVCYQRY 126
 QY 354 KEFGNKEISQGLSLFLALMHGILHSGYLVCQMEFLIVEROARLIQESPTLSKLA 413
 DB 127 PWYPTV-----LTFILSALMHGVPCYPTFLGILVTLAAR-AVRNNRHYTLSSRA 180
 QY 414 TVLQPFYVYVQQTIMLFMGYSNTAPCLFTWDMKLVKYSIYFLGHIFFLSLFIPIYH 473
 DB 181 AVYDAGTMAVTQ---LAVSYTVAPFVMLAVEPTISLYKSMYFYLIHISLITILFPMKP 236
 QY 474 KAMVPRK 480
 DB 237 QAHTORR 243

Search completed: November 21, 2003, 14:14:42
 Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:12:23 ; Search time 21 Seconds
(without alignments)
2230.198 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 2580
Sequence: 1 MASSAECDEGTVALACVLQ.....ILPYHKAWPRKXKXKME 487

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	749.5	29.1	486	2 T15281	hypothetical prote
2	407	15.8	465	2 G96655	unknown protein, 2
3	406	15.7	509	2 T139542	hypothetical prote
4	406	15.7	619	2 S67067	probable membrane
5	383.5	14.9	480	2 T15839	hypothetical prote
6	297	11.5	540	2 T19097	hypothetical prote
7	244.5	9.5	480	2 T20899	hypothetical prote
8	158.5	6.1	366	2 T27912	hypothetical prote
9	147	5.7	135	2 T05909	membrane protein y
10	116.5	4.5	477	2 AB2313	hypothetical prote
11	116.5	4.5	736	2 C69451	cationic amino aci
12	115	4.5	392	2 B96610	hypothetical prote
13	114.5	4.4	609	2 S65208	probable membrane
14	114	4.4	614	2 B84949	NADH2 dehydrogenas
15	113.5	4.4	547	2 T13437	NADH2 dehydrogenas
16	111.5	4.3	419	2 E90446	permease [imported
17	111	4.3	392	2 A65086	probable hydrogena
18	110.5	4.3	355	2 T13871	NADH2 dehydrogenas
19	110.5	4.3	355	2 T13872	NADH2 dehydrogenas
20	110.5	4.3	355	2 T13873	NADH2 dehydrogenas
21	110.5	4.3	355	2 T13869	NADH2 dehydrogenas
22	110.5	4.3	355	2 T13870	NADH2 dehydrogenas
23	110	4.3	396	2 S49592	cytochrome b-like
24	109.5	4.2	355	2 T13873	NADH2 dehydrogenas
25	109.5	4.2	355	2 T13911	NADH2 dehydrogenas
26	109.5	4.2	355	2 T13912	NADH2 dehydrogenas
27	109.5	4.2	355	2 T13912	NADH2 dehydrogenas
28	109.5	4.2	355	2 T14092	NADH2 dehydrogenas
29	109	4.2	392	2 H91113	hypothetical prote

30	109	4.2	392	2 H85958	hypothetical prote
31	108.5	4.2	355	2 T13986	NADH2 dehydrogenas
32	107.5	4.2	355	2 T13865	NADH2 dehydrogenas
33	107.5	4.2	355	2 T13909	NADH2 dehydrogenas
34	107.5	4.2	355	2 T13910	NADH2 dehydrogenas
35	107.5	4.2	355	2 T13908	NADH2 dehydrogenas
36	107	4.1	352	2 T11734	NADH2 dehydrogenas
37	107	4.1	392	2 AD0885	probable hydrogena
38	107	4.1	619	2 T11314	NADH2 dehydrogenas
39	107	4.1	724	2 T19601	hypothetical prote
40	106.5	4.1	355	2 T13919	NADH2 dehydrogenas
41	106.5	4.1	355	2 T13987	NADH2 dehydrogenas
42	106.5	4.1	355	2 T13984	NADH2 dehydrogenas
43	106.5	4.1	355	2 T13985	NADH2 dehydrogenas
44	106.5	4.1	537	2 T41664	amino acid permeas
45	106	4.1	355	2 T11739	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1
T15281
hypothetical protein R155.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T15281
R/Geisel, C.; Mameley, P.; Kramer, J.
Submitted to the EMBL Data Library, May 1997
A/Description: The sequence of C. elegans cosmid R155.
A/Reference number: Z18321
A/Accession: T15281
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-486 <GEI>
A/Cross-References: EMBL:AF003390; NID:G2088866; PID:G2088866; PIDN:AA854271.1; GSPDB:GN
A/Experimental source: strain Bristol N2, clone R155
C/Genetics:
A/Gene: CESP.R155.1
A/Map position: 3
A/Introns: 27/1; 127/2; 171/3; 203/2; 289/2; 324/2; 337/2; 439/3
Query Match 29.1%; Score 749.5; DB 2; Length 486;
Best Local Similarity 32.8%; Pred. No. 5.7e-52;
Matches 157; Conservative 95; Mismatches 203; Indels 23; Gaps 4;
QY 31 LATSLSAQSEALRLIISIFLGYPPALFRRHYLFYKETYLLHLPHFTGSLIAFNFGNQL 90
DB 7 LSEVTSASEDALRLISLVLAGYPLAVVHRTFYVKPAQHOLHFLVIGLSLIMFNCSSV 66
QY 91 YHSLLCTVQLFLRLRMGRITTVLTFFCFQMAVLLAGYVYTAAGNDIKMTMPHCVLTL 150
DB 67 IHPILSLFGAFFTTNFAGTDASTYAAHIVFLGHLGVMFHEVDYDITMTTTPCCIMTL 126
QY 151 KLIGLAVDFPDGQKQNSLSEQOKVIRGVPSLLEVAGSFYFAGFLVGFQSFNNHYMK 210
DB 127 RFLGLVWDVYDAGQKPRHLKPDQKLTISDKPGLLEIAAGLFGQTLVGFQFLSLKFRS 186
QY 211 LVQGLIDIDPKLINS-IIPALRSLGLFLYLVGYTLLSPHTEIDYLLTDDYDNPWFWR 269
DB 187 FVNDWDVSDGQPPKSAFLPSIGRFLAGCTYVMVHLQMGQFIMPQYFNSDAVNNLSFPMR 246
QY 270 CMVWLWLGKFLVLYVYVCMVLVBGVCLTGLGFGFGEKRAKADACANMKWMLFENNPR 329
DB 247 WSWVTLWFLRLMYKCMMLITBGSATLSGLGHGKABEGNDRDVGRDLITIKWETGHD 306
QY 330 FTGTIASFNINTAMVAVR-----YIFKRLKFLGNKELSGQLSLFLALMHLG 376
DB 307 YNSVSEFNGCTINFPAKQNFALFNIPATKIHHRRLRWNNKLASAVITLSYALIMHG 366
QY 377 HSGYLVCQMFLLIVVERDPAARLIQSPSLSKLAATVLAQPFYLVQOITHLWFGYSM 436
DB 367 HLGYFLFGLVGLGCVQKQNDLYALIKRTIPGMSEAIKSPISRPFIWIFGKLTISYSMGFAF 426

QY 437 TAFCLFTWDMKMLKVKYSIYFLGHIFP-----LSLFLIP-YIHKAMVPRKELKK 485
 DB 427 LMFGLIKTKYKVIIGPKYSIYFIFGIYFIYFWPLIMVILNVLNRPKKAALAEKPEEVKK 484

RESULT 2

G96655
 unknown protein, 29405-27288 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: G96655
 R:Theologos, A.; Becker, J.R.; Palm, C.D.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Malt, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A66141; MIMD:21016719; PMID:11130712
 A:Accession: G96655
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-465 <STO>
 A:Cross-references: GB:AE005173; NID:g6598844; PIDN:AAPI8699.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F16M19.14
 A:Map position: 1

Query Match 15.8%; Score 407; DB 2; Length 465;
 Best Local Similarity 25.4%; Pred. No. 9.8e-25;
 Matches 123; Conservative 90; Mismatches 227; Indels 44; Gaps 12;

QY 26 LSLINKLATSLSGSEALRLIISIFLGYPALFYRHYLHFTFTGLSIAYFN 85
 DB 4 LDMNSMAISIGVAVLRFLLCFVATIPISFLMR---FIPSLKNIYAAAGAFSLYS 60
 QY 86 FGNOLYHSLLCIVLOFLILRLMGRTTAV-----LTFQOMAYILAGYTTATGN-- 137
 DB 61 FG-----FSSNLHFLVMTIGYASMAIYRPLSGFTFFLGFAYILGCHVFMGDMK 113
 QY 138 --DIKWTMPCVLTLLKLGAVDPDGKDONSLSSQOKYAIRGVSFLLEVAGSYFPG 195
 DB 114 EGGISTGTALMVLTLKVISCSININDGMKEGRLPAQKKRLIQPISLETFTGLCCG 173
 QY 196 AFLVGPQFSNMHYMKLVQGE---LIDIPKINSIIPALKRLSLGLFVLVGYTLSPHIT 252
 DB 174 SHFAGPVPFEMKDYLEWTEKGIMAVSEKGRPSVGMIRAVFOAICMALVLYLVPDP 233
 QY 253 EYVLLTEDVDNHPFMRCHYMLIMGKFVLYKVTCTLVTEGVCILTGLGFNGF--EEKK 310
 DB 234 LTRFTEPVYQEWGFLKRFQYQYAGFTARKYFIWISSEASIIISGLQFSGMTDETQK 293
 QY 311 AKMDACANNKVMFLFTNPRFTGTIASFNINTNAAVARYIFKRLKFLNGK-----LSQG 364
 DB 294 AKMDAKAVDIIIGVELAKSAVOIPLFWMNQVSTWLNHYERIVKRGKAGFOLLATQT 353
 QY 365 LSLFLALMGLHSGYLVCQMEFLIYVEROARLIQ--SPTLSKLAITVLOPFYVLV 423
 DB 354 VS-----AVMHGLYPGIIIFVQSAIMIDSKAIYRQQAIPKRMALRNVLVLINFLYV 409
 QY 424 QOTIMLFMGYSMTAFCLFTWDMKMLKVKYSIYFLGHIFLSLFLIYIHKAMVPRKEL 483
 DB 410 -----VVLNYSVGFMVLSLHETVLAFAKSVYIIGTIVIAVL-LISYL-VVKKVPKPT 461
 QY 484 KQME 487
 DB 462 KKEE 465

RESULT 3

hypothetical protein SPBC16A3.10 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: J39542
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, February 1998
 A:Reference number: J21862
 A:Accession: J39542
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-509 <WOO>
 A:Cross-references: EMBL:AL021748; PIDN:CAA16861.1; GSPDB:GN00067; SPDB:SPBC16A3.10
 C:Genetics:
 A:Gene: SPDB:SPBC16A3.10
 A:Map position: 2

Query Match 15.7%; Score 406; DB 2; Length 509;
 Best Local Similarity 25.1%; Pred. No. 1.3e-24;
 Matches 121; Conservative 86; Mismatches 206; Indels 70; Gaps 13;

QY 25 ELSINKLATSLSGSEALRLIISIFLGYPALFYRHY-----LFYKEYTL 69
 DB 6 DIPFVFSSFLGVHPDQKLFLCFLSAVFPAGILKRLSPAWIRNLFSISGLF----YL 61
 QY 70 IHLFTFTGLSIAYFN-----FGNOLYHSLLCIVLOFLILRLMGRTTAVLTFQOMAY 124
 DB 62 IGVMHLVGVVLVLFDAFTYFAAFYRSSMPWIIPIV--ILGHTFSSHVIRYIY---- 115
 QY 125 LLAGYVYATGNYDICKMTMPCVLTLLKLGAVDPDGKDONSLSSQOKYAIRGVSFL 184
 DB 116 -----PSNTDI--TASQVLCMKLTARMSVYDRLSSLSSTYQKRALKRTVIT 165
 QY 185 LEVAGFSYGAFLVGPQFSNMHYMKLVQ---GELIDIPK--IPNSIIPALKRLSLGL 238
 DB 166 LYFLGVFPFSLVLAGPAFDVYDERFTLSMPKRLADPYEKQITPHSLPALGCMWGL 225
 QY 239 FYLVGYTLSPHITEDYLDNHPFMRCHYMLIMGKFVLYKVTCTLVTEGVCILT 298
 DB 226 LMLIFITGSSIYPLKFLTPKFASSPILLYGYVCIATFAARMKYGAWELSDACILS 285
 QY 299 GLGFNGFEKQAKMDACANNKVMFLFTNPRFTGTIASFNINTNAAVARYIFKRLKFLGN 358
 DB 286 GIGVGLDSSKHPKRDYKVIDPIKFEPADNIKCALEAMNNTNKMNLNVIYLRVAKKXK 345
 QY 359 KE-LSQGLSLFL-ALMHGLHSGYLVCQMEFLIYVEROARLIQ-----ESPTLS 408
 DB 346 RPFKSTLSTFTVSAMMHGVSAGYLLTFVSAFIQTVAKYTRRHVRPFLKPDMEITP--- 402
 QY 409 KLAITVLOPF---YVLVQOTIMLFMGYSMTAFCLFTWDMKMLKVKYSIYFLGHIFLSL 465
 DB 403 -----GPFKRVYDVIGVATNLSLYLIIISFLNLKESIHVMKELYFIIVHILIA 454
 QY 466 LFI 468
 DB 455 LAV 457

RESULT 4

probable membrane protein YOR175c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein OJ635
 C:Species: Saccharomyces cerevisiae
 C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
 C:Accession: S67067; S67063
 R:Hughes, B.; Pohl, T.M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67067
 A:Accession: S67067
 A:Molecule type: DNA
 A:Residues: 1-619 <HUG>

A:Cross-references: EMBL:Z75083; NID:g1420424; PID:e253056; PID:g1420425; MIPS:YOR175c
 A:Experimental source: strain S288C
 R:Borderline, R.; Camases, A.; Madan, A.; Martin, R.P.; Poch, O.; Tarasov, I.A.; Winsc
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67032
 A:Accession: S67063
 A:Molecule type: DNA
 A:Residues: 270-619 <BOR>
 A:Cross-references: EMBL:Z75083; MIPS:YOR175c
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0005701
 A:Map position: 15K
 C:Keywords: transmembrane protein
 F:53-69/Domain: transmembrane #status predicted <TM1>
 F:461-477/Domain: transmembrane #status predicted <TM2>

Query Match 15.7%; Score 406; DB 2; Length 619;
 Best Local Similarity 26.2%; Pred. No. 1.6e-24;
 Matches 133; Conservative 79; Mismatches 238; Indels 58; Gaps 17;

OY 28 LNKATSLGASEQRLRLISIFLGYPALFVRH-----LFYETVLI--HLFHTPTGLSI 81
 DB 9 LTKITITNGIDSFTRVAILCLGSPFLNALIKRIPKRIKGLCCPTISMSPYLF----- 63
 OY 82 AYFNGNOLYSLICIVLPILRLMGRITVAVLTTCFOMAYLLAGYVYATGNYDIKW 141
 DB 64 GVLNVLVSFRTLLFTMTFTYLSIRFYRSKMPHL-NFNFVGHGLAINHHQFLNEQOT 122
 OY 142 TM---PHCVTLTKLIGLAVDYFDG---KDQNSLSSEQKVAIRGVPSLLEVAGSFY 193
 DB 123 TVDITSSQVVLAMKLTSFAMSYDGSCTSESDFKDLTEHQKSRVAGHPPLKFLAVAF 182
 OY 194 YGAFVGVQFPMNHNKLVQBELI-DIP-----GKIPNIIIPALKRSLGL 238
 DB 183 YSTLTGSPFYADDSWLNCEMFRLDPESKKPMRHHGERRQIPKNGKALMWVQGL 242
 OY 239 FYLVGYTLSPHTEDEYLTED-YDNHPEFRCMYLMIGKFLVLYKYVTCMLVTEGVCIL 297
 DB 243 AMMLISTIGMGHFPKYVLYDKGFPTRSPFIRIHLPLFLGFIHRKYVAAMTISGSCIL 302
 OY 298 TGLGNGEERK-GKAKWDACANMKWLFETNPRFTGTIASFNINTANVAARIYFRKLFL 356
 DB 303 CGLGNGYDSKTQKIRMDRVNRIDIVETAONTREMEANMNNTNKKLVSYLVRTKK 362
 OY 357 GNKELSQGLSLFL--ALMHGHSGLVCFQMEFLIVVERQAALIOES--PLTSKLA 412
 DB 363 GKRPGRSTLFTPLTSAPWHGTRPGVYLTFATGALY---QTCGKIRNRNRPFLFLREDG 418
 OY 413 ITVL--QPFYLVQOTIHWLFMGYSMTAFCLFTWDKMLKVKYSIYFLGHIF--FLSLFL 468
 DB 419 VTPLESKITVLDVGLYAKLAFLAGVAVQPFIIIDLKPSLAWGVSIFYVYIIIVASFLFLR 478
 OY 469 LPY-----HKAVNPRK--EKLKME 487
 DB 479 GPYAKVTEFESKQPKEIFIRKQKLE 506

RESULT 5
 T15839
 hypothetical protein C54G7.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 15-Sep-2000
 C:Accession: T15839
 R:Du, Z.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid C54G7.
 A:Reference number: Z18416
 A:Accession: T15839
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-480 <DUG>
 A:Cross-references: EMBL:U040410; NID:g1065453; PID:g1065454; PIDN:AAA81391.1; CESP:C54G7

C:Genetics:
 A:Gene: CESP:C54G7.2
 A:Introns: 27/3; 76/2; 102/2; 185/3; 283/1; 384/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C08F8.4

Query Match 14.9%; Score 383.5; DB 2; Length 480;
 Best Local Similarity 24.1%; Pred. No. 7.5e-23;
 Matches 112; Conservative 98; Mismatches 221; Indels 33; Gaps 13;

OY 28 LNKATSLGASEQRLRLISIFLGYPALFVRHLPFKET--YLHLFHTPTGLSIAYFN 85
 DB 14 LEPLAETIGLEADRVNVLSTFACFGISYVRKTCSLQVNRQVTVPAVAGIYVFC 73
 OY 86 FGNOLYSLICIVLPILRLMGRITVAVLTTCFOMAYLLAGYVYATGNYDIKW 144
 DB 74 FGRAIKHLNALAGSVAIMYAPPTQHKVILFWSMGLFPIHCYRNALLESYLDVTCF 133
 OY 145 HCVLTLLKIGLAVDYFDG-KDQNSLSSEQKVAIRGVPSLLEVAGSFYGAFLVGPQF 203
 DB 134 IMVAVERATWMAFNLKDGKADQSKLTEEQKRESLKEIPSLLEFMSFENFQVLTGPAN 193
 OY 204 SMNHYMLVQGE--LIDIPKIPNSIIPALKRSLGLFYL-VGYTLSPHTEDEYLTED 260
 DB 194 NYDYIKFLDBKHLVADKHGLPSPGTAAAMKRYQSIFFLAIVTLTGKRYVED-VGTOP 252
 OY 261 YDNHPE--WFRCMYLMIGKFLVLYKYVTCMLVTEGVCILTGLFNGEERKAKWDACAN 318
 DB 253 YFALPFPQMFWMWITTF--FIRCAVYFANVPAIACMSGFSGYDKENAEKLCITN 310
 OY 319 MKWLFETNPRFTGTIASFNINTANVAARIYFKRLKFLGNKELSQGLSLFLALMHGHS 378
 DB 311 VLPVQVEMAQSLKETLDQMNIOGTGFWLRKGYER---APKSIPTVATVYLSAVWHGYSI 366
 OY 379 GYLVCQMEFLIVVERQAALIOES-----PLTSKLAATTVQPFYLVQOTIHWLFM 432
 DB 367 GYVMAFTCGFTV---AAQTFRRSMKWRFLDHNKFAVDI--FSFIISK---IAL 415
 OY 433 GYSMTAFCLFTWDKMLKVKYSIYFLGHIFPLSLFLIYPIYIKAM 476
 DB 416 AYATYSFTWMTMYPAYDVLKRVYFIPIHVAIVAFIPALPKFVKPL 459

RESULT 6
 T19097
 hypothetical protein C08F8.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
 C:Accession: T19097
 R:Harris, B.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19073
 A:Accession: T19097
 A:Map position: 4
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-540 <WIL>
 A:Cross-references: EMBL:Z73103; PIDN:CAA97423.1; GSPDB:GN00022; CESP:C08F8.4
 A:Experimental source: clone C08F8
 C:Genetics:
 A:Gene: CESP:C08F8.4
 A:Map position: 4
 A:Introns: 35/3; 82/2; 108/2; 146/3; 210/3; 309/1; 344/3; 411/3; 462/2; 515/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C08F8.4

Query Match 11.5%; Score 297; DB 2; Length 540;
 Best Local Similarity 25.0%; Pred. No. 6.6e-16;
 Matches 114; Conservative 71; Mismatches 171; Indels 100; Gaps 22;

OY 75 TPTGSLAYENFGNOLYSLICIVLPILRLMGRITVAV-----LTFPCOMAYTL 126
 DB 69 TSVGLIFTFYCYGNAIAH-----LFINGFSYLLMVSVPQHNKSVFAPAMGYLV 119
 OY 127 A-----GYVYATGNYDIKMTNPHCVTLTKLIGLAVDYFDG-KDQNSLSSEQKVAIRGV 181

[illegible]

```

RESULT 7
T20899
hypothetical protein Fl4F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20899
R:Kershaw, J.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19343
A:Accession: T20899
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-480 <MIL>
A:Cross-references: EMBL_Z49937; PIDN:CAA0165.1; GSPDB:GND0028; CESP:Fl4F3.3
A:Experimental source: clone Fl4F3
A:Gene: CESP:Fl4F3.3
A:Map position: X
A:Introns: 25/1; 79/1; 102/2; 145/3; 192/1; 241/3; 273/2; 331/3; 385/2; 420/2

Query Match          9.5%; Score 244.5; DB 2; Length 480;
Best Local Similarity 21.3%; Pred. No. 8.7e-12;
Matches    96; Conservative   91; Mismatches 184; Indels   79; Gaps 16;

QY      GLSIAYFNFGQLVSL---LCIYLQGLIRLMGRTTAVLTTPCFOMATVLGGYYTA 133
       |||::||::||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      70 GFANSLFIIGPKIVISLGICISIAISIQ-LANKSKSTPIYVELITTYTM-EVRFAHYILP 127
       ::||::||::||::||::||::||::||::||::||::||::||::||::||:

QY      134 TGNYDIKMTWPHCVTLTKLIGLAVDYFDG---GKDONSASSBOOKYAIRGVPSLLEVAGF 190
       :||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      128 VNEVASHTNVLIQTILTRIGITPEENDAMVHKSDEN-----PTRKYVTLPETLEKRAY 182
       ::||::||::||::||::||::||::||::||::||::||::||::||::||:

QY      191 SYFYCAFLVVGQFSNNHMKLVQGLLDIPCKIPNSIIPALK--RLSLGIPLYVGYTLL 247
       |||::||::||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      183 FYHFQGLFTGPYYTQQ-----MLDISONPILKSMPTLEVKSRFVRLIMSVAPVYIT 234
       |||::||::||::|||::|||::|||::|||::|||::|||::|||::|||::

QY      248 SPHIREDYLLIEDYDNHDFPFRCMMVLIMGKVLYK--YVMCMVTEGVCLITGLG---- 301
       :||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      235 NHYPLDLILRSDALWEVFRTLRVYVAL--IFVFKTVYSMAWAIESICVILGIGIYP 292
       :||::||::||::||::||::||::||::||::||::||::||::||::||:

QY      302 -----FNQFE-----EKGRAKMDACANNKWLFETNPRTGTIASFNINTNW 344
       :||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      293 ASNPKIIMGPTDLNAFDLKTRENIEMSSDAIVNLDPKEVRSDFRGDKAMNRSVQTW 352
       :||::||::||::||::||::||::||::||::||::||::||::||::||:

QY      345 VARYIFKRLKFLGNKELSGSLFLALMHGHSGYIVCFQMELIIVVERKQARLIOES 404
       :||::||::||::||::||::||::||::||::||::||::||::||::||:

```

```

Db      355 LALVHSRKNKRV-----TMTLSAVMHGTAGYFMSPSVAMCMLLEDVIFKLV---404
Qy      405 PTSKMAATVUQPFYLVQQTIIHLMFNGYSM--TAECLFTWDMKLVKYSIYELGHI--460
Db      405 -----PVDTEGVAPKPFRLIYTHITTRCGFEELATGFLD-----KNAVDFHFMWS450
Qy      461 ---FPLSLFLPYIHKAVPRKEKTKKME487
Db      451 SIYYWLPCLCIPEYIYSAKISPKKQAKSE480

```

RESULT 8
T27912
hypothetical protein ZK550.1 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 06-Oct-2000
C|Accession: T27912
R|Basham, V.
submitted to the EMBL Data Library, November 1996
A|Reference number: Z20438
A|Accession: T27912
A|Status: preliminary; translated from GB/EMBL/DDbj
A|Molecule type: DNA
A|Residues: 1-366 <WtL>
A|Cross-references: EMBL:Z82287; PIDB:CA05316.1; GSPDB:GN00022; CESP:ZK550.1.1
A|Experimental source: clone ZK550
C|Genetics:
A|Gene: CESP:ZK550.1
A|Map position: 4
A|Intons: 35/3; 82/2; 108/2; 146/3; 191/3; 266/3
C|Superfamily: *Caenorhabditis elegans* hypothetical protein C08F8.4

```

Query March 6.1% Score 158.5; DB 2; Length 366;
Best Local Similarity 21.6%; Pred. No. 4.5e-05;
Match 78; Conservative 57; Mismatches 11; Indels 15; Gaps 16

Dy 78 GLSIAYENFGNQLVHSLLCIVLQFLILRLMGRITLAVL-----"TTFCFOMAYL-LAGYL 130
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 72 GPILFYFCYGNELIAHFPIINGFGSYILM-----ISVLPRKVHKAVFSPAMAYLFLVHFY 124
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 131 ---YTATCNVYDIKMTMPHCVLTLKILGLAVDFDCK-DONSLSSEQKAIKVPSLIE 186
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 125 RWIWOETFNILFTSGSMVAVAVKITLSSAIT-DGMRDLDALNSGQRDAVNEIPSLID 182
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 187 VAGSFYFAGLTVGQPSMNH-----MKLV-QGELIDIPGKIIPKSIIPALK-RLSL 236
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 183 FASITFAFQSYITGF---TNHYSMSAVLDLTLPKFKRTGRPFPGSTSTVEKFKVALTL 239
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 237 GLFYLVGYTLTSLPHITEDYLLTEDVDNHPFWFRCMYLMIMQKFVLKYVTCWLVTGVC 296
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 240 SSFCSALYSLPIPLTKYPTISE-----YNLL----- 266
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 297 LTGAGFNGFEKGAKKADCANMKMYLFETNPRFGTITIASFNININNAVARIIPKRLFL 356
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 267 -----STNCNVW-----TDALRLRIYERV--- 286
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 357 GNKELSQGLSLFL-ALMHGLSHGYLCFQMEFLIVVEROARLIQESPTSLKTAAT 414
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 287 ---EDPRKVIAYVYGVAMHGLAVRYISFLTSAIFTLAKVGSITLFO---AICMAANT 339
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 415 V 415
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 340 V 340

RESULT 9
T05909
membrane protein YOR175C homolog - barley (fragment)
C:Species: Hordeum vulgare (barley)
C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 08-Oct-1999
C:Accession: T05909
R:Hees, W.R.; Goliz, R.R.; Boerner, T.
Plant Sci. 133, 191-201, 1998

```



```

0Y 26 L$INKLATSLSGSEQAL---RLIISFLGYPFALFRHYLLEYETYLHLPHTGSLTA 82
Db 68 L$INFEKIDFGFLLDGL$LSMLFVITGVGLLHIFBSWYMRKYEGQ--SRFPAYINLFTA 125
QY 83 YFNFGNOLYHSLLCIVLOFLILRLMGRTTAAVLTPFCOMAYLLAGYUYTATGNDIKWT 142
Db 126 SM-----SYLVADNLFPMYLGEGVSV-----CSYLLIFGYTELKNM--LCA 167
QY 143 MPHCVLT-----LKLIGLAVDPFDGKDONSLSSEQOKYAIKGVPSLLEVAGHSYRYGAF 197
Db 168 FKAFPLTRVSDVFLMIGMFLYI---RENSRFPQKIFK---LSSFLVNEANYI--- 215
QY 198 LVGPOFSMNHMYKLVQEGELIDIPKIPN$ILPALKSL$GLFPLVGYTLLSPHITEDYLL 257
Db 216 -----DYITLFL--L$GVYIK--SAQPLQOTMLSDA---MVEPTV$NLIHATWV 259
QY 258 TEDYNNH$PFRCMYMLMGKFLVLYKYVT$CMLTVBSVCILITGLGFNGFBEK$AKADACA 317
Db 260 TA-----GVYLIARTH$FLITP$ILVGL----- 285
QY 318 NMK$VLFETPNRPTGTIAS$FNINTAN$VARYFKRLK$FNGKEL$OGSL$FLAL-----W 373
Db 286 -----IGTITLIV$SISALVOKDI--KR--LAYSMSQ--IGMFLALGV$KXN 328
QY 374 HGLHSGYLVCFQMEFLIVIVERQ$ARL$IQ$ESPL$K$AIAITVLOP$YVYLVOQT$HMLFNG 433
Db 329 SAATLHLLIMHAI$FKALL$F$AGSL$IL$CKN$ENKI$F$MGL$IRQL$PFLY-----ISP$VCG 383
QY 434 Y$MTAF$CL$ETMD$KMLKVY$SIYFL$HIF 462
Db 384 ASLVSFPLIT-----AGFY$K$G$ILF 404

```

```

RESULT 15
T33437
hypochemical protein ZC190.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33437
R:Nelson, J.; Wohlmann, P.
A:Description: The EMBL Data Library, July, 1998
A:Reference number: Z13344
A:Accession: T33437
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-547 <NLS>
A:Cross-references: EMBL:AF078788; PDB:1AC25966.1; GSPDB:GN00023; CESP:ZC190.6
A:Experimental source: Strain Bristol N2; clone ZC190
A:Genetic:
A:Gene: CESP:ZC190.6
A:Map position: 5
A:Introns: 40/2; 105/1; 165/3; 254/3; 272/2; 338/1; 387/1; 406/3; 433/3; 507/3

```

Query Match 4.4%; Score 113.5; DB 2; Length 547;
Best Local Similarity 19.2%; Pred. No. 0.27;
Matches 95; Conservative 113; Mismatches 184; Indels 103; Gaps 26;

```

OY 62 FFYKETYLIHLEHTTGTGSLAFNGNQLYHSLCLVLOFLIRLM---GRITAVLTTP 118
Db 45 LFSVSIYLLDISH---GLTVFEELFEKQTYHIIHFVSVGITWRFLHEFGYSHQSLSLF 100
OY 119 CFQMAVYLAGYYTATNGYDIKMTBPHCVLTILKLGANDVYEDGGDQNSLSSEQKVAI 178
Db 101 LFSIERLLASLOPTRYRSEKIRIWIPIFSLGSIISIGFSY-----TVHGMQMFSP 152
OY 179 RGVPSLELVAGSFYVYGAFLVQPSMNMHYMLV-----QGLIDIDGKIPNLSITPVL 231
Db 153 TVAMVLVDISSCIF---LIIMQKSLSHYRSYGVGLTNORFQSDV-YTWSRHLIPAV 207
OY 232 KRSLSGLF-YLVGVTLLSPHTED---YLLTEDVDN-----HPFW-----PQC 270

```

Db 208 LLASLKSPKICITIIWLLITSNTFENQQAIIWIIA--YNNILINIVAVILPFLVQSTVNFOL 265

Qy 271 MYWLLIGKRVLYKVVTCMLVTBGCIVLLTGIGFENPE-----EKGAKKADACAKNMVLMFEF 326

Db 266 IHHAKKASPELLSPSSLPFLA---LTH-GEQIFELLIGADANAIDYTSNRMLVMSL 319

Qy 327 NPFRTGIASPNIINTANAVARIYFRKLK-FLGKNKELSGSLFLALMHGLSGYL-VC 363

Db 320 LHFEAFALSPSIQMFSDSKHTLTLERYITTAQOTIQSVRNOVLSTLGVIIISILSMC 379

Qy 384 EQ-----MELLV-----IYERQARLLIOESPILLS-----KLAIVLQGFY-----YL 422

Db 380 YAOIHLVDFHVIATITMNLVDMSLVLIIISQYSIONRKTGINSLEKRFQISDVI 439

Qy 423 VQGT-----IHWLFGYSMTAFCLFTWD--KMLVYKYSIYF--LGHIFELSLFT 468

Db 440 WFOAIIIPVACIAFLIHFFV--IVWMLSTLSENSRYQMTPSHYFENITSVFICVPL 496

Qy 469 LPYIHKAMVPRKEXL 483

Db 497 VYIKAKRI--KTKL 509

Search completed: November 21, 2003, 14:16:37
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:03:03 ; Search time 17 Seconds

(without alignments)
1347.177 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 2580
Sequence: 1 MASSAEGDEGTVALAGVLQ.....ILPYIKAMVPRKEKLMKE 487

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	121	4.7	615	1	NNDL_BUCAP
2	114	4.4	614	1	NNDL_BUCAI
3	109	4.2	588	1	YV71_SCHPO
4	107	4.1	380	1	CYB_GARNE
5	107	4.1	380	1	CYB_PELGA
6	107	4.1	392	1	HYBB_ECOLI
7	106.5	4.1	537	1	YCUB_SCHPO
8	106	4.1	512	1	COXI_LOCM1
9	105	4.1	380	1	CYB_PREHY
10	105	4.1	521	1	COXI_APIIL
11	105	4.1	545	1	NUSM_ALBGO
12	105	4.1	654	1	NUSM_RHIST
13	104.5	4.1	380	1	CYB_BRALA
14	104.5	4.1	395	1	DLTB_BACSV
15	104	4.0	379	1	CYB_TINMA
16	103.5	4.0	1060	1	NKCL_MANSE
17	102.5	4.0	461	1	MVIN_HELPY
18	102.5	4.0	514	1	COXI_PELSU
19	102.5	4.0	546	1	Y61B_MYCPN
20	102	4.0	379	1	CYB_URSAM
21	102	4.0	498	1	TLCX_RICPR
22	102	4.0	513	1	COXI_MACRO
23	102	4.0	527	1	NUSM_ACACA
24	101.5	3.9	549	1	COXI_LETRA
25	101.5	3.9	686	1	NUSM_SCHCO
26	101	3.9	380	1	CYB_POLBI
27	100	3.9	380	1	CYB_PAVMU
28	100	3.9	380	1	CYB_PUPTU
29	100	3.9	501	1	ANKH_BRARE
30	100	3.9	919	1	YKOS_YEAST
31	99	3.8	380	1	CYB_CALLE
32	99	3.8	380	1	CYB_LACV
33	99	3.8	380	1	CYB_LOPGA

ALIGNMENTS

RESULT 1	ID	NNDL_BUCAP	STANDARD	PRT	615 AA.
AC	08K9X7				
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	NADH-quinone oxidoreductase chain L (EC 1.6.99.5) (NADH dehydrogenase I, chain L) (NDH-1, chain L).				
DE	NNDL OR BUSG157.				
OS	Buchnera aphidicola (subsp. Schizaphis graminum).				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Buchnera.				
OX	NCBI_TaxID=98794;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22084549; PubMed=12089438;				
RA	Tamas I., Klasson L., Canbeck B., Naeslund A.K., Eriksson A.-S.,				
RA	Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;				
RT	"50 million years of genomic stasis in endosymbiotic bacteria.";				
RL	Science 296:2376-2379(2002).				
CC	-1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-sulfur (Fe-S) centers, to quinones in the respiratory chain.				
CC	Couples the redox reaction to proton translocation (for every two electrons transferred, four hydrogen ions are translocated across the cytoplasmic membrane), and thus conserves the redox energy in a proton gradient (By similarity).				
CC	-1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.				
CC	-1- SUBUNIT: Composed of 13 different subunits. Subunits nuoH, H, J, K, L, M, N constitute the membrane sector of the complex (By similarity).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).				
CC	-1- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE OF CHLOROPLASTS OR MITOCHONDRIA.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL; AE014091; AAM67725.1; -				
CC	InterPro; IPR003918; NADHnb_oxred4.				
CC	InterPro; IPR003916; NADHnb_oxred5.				
CC	InterPro; IPR001750; Oxidored_q1.				
CC	InterPro; IPR001516; Oxidored_q1_N.				
CC	Pfam; PF00361; Oxidored_q1; 1.				
CC	Pfam; PF00662; Oxidored_q1_N; 1.				
CC	PRINTS; PR01434; NADHGNASE5.				
CC	PRINTS; PR01437; NIOXDRDTS4.				
CC	Oxidoreductase; NAD; Quinone; Transmembrane; Complete proteome.				
FT	TRANSMEM 1 21				
FT	TRANSMEM 32 52				
FT	TRANSMEM 79 99				
FT	POTENTIAL.				
FT	POTENTIAL.				
FT	POTENTIAL.				

```

FT TRANSMEM 113 133 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 279 299 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 372 392 POTENTIAL.
FT TRANSMEM 410 430 POTENTIAL.
FT TRANSMEM 457 477 POTENTIAL.
FT TRANSMEM 491 511 POTENTIAL.
FT TRANSMEM 536 556 POTENTIAL.
FT TRANSMEM 594 614 POTENTIAL.
SQ SEQUENCE 615 AA; 70511 MW; 8873DB384AB7BE1 CRC64;

```

Query Match 4.7%; Score 121; DB 1; Length 615;

Best Local Similarity 22.0%; Pred. No. 0.14; Indels 182; Gaps 32;

Matches 124; Conservative 67; Mismatches 191; Indels 182; Gaps 32;

```

QY 20 QSGFQELSLNKLATSL-----GASEQALRLIISIFLGYPPALFYRHYLFYKETYLIHLPH 74
DB 62 QKLFMSISINELDIDCSLHLDGSLSMILMIGI--GLIHIFSTWYMDKGGY--SRPF 117
QY 75 TTGTGSIANFNQNLHSLCTIVLOFLIRLMGRTTAVLTTCFQMAVLLAGYYTAT 134
DB 118 AYTNIPLIASM-----SLVLADNLFEMYL-GWEIVSI-----CYLLIGFYTKT 161
QY 135 GNYDIKMTWPHCVLTCLKGLAVDPDGKDNSSLEQOKYAIRGVSPLLEVAGSFYF 194
DB 162 NN-----TSCALKAFTFTRISDV-----LIISMFLYKGTGTFNQEIKLSNPL 207
QY 195 GAFVGPQPSNMHWML-----VOGELIDIP-----GKIP-NSIIPALKRLSLGL 238
DB 208 N--VEDCPDLN-VLTCLLGLWGKSAQLPLHTWLSDMVGPPTPSALIHATWTAAGV 263
QY 239 FVLVGYT---LISPHITEDYL-----TEDYDHPF 266
DB 264 -YLIRATHFLFLTPKIL--YLISLIGITITPISFSPALVODIKRILAVTWSQIGYMF 320
QY 267 -----W-----PRCMVYLIMGKFL-----YKYVTGW 288
DB 321 LALGVKAMPAIYVHLIYHAIFFKALFLSSGSLSCNNENKFNLSKVSCKEPLVVS-F 379
QY 289 LVTEGVCIITGLGFNGFEKGRKAKMDACAMKVMFETNPRFTGTIASFNINTANAVARY 348
DB 380 LVGASLVSFPLITSGFYSKGNILFSLVDGYFNLF-----LIGLCSF--LTSITFFM 432
QY 349 IF-----KRLKFL-GKKEISQGLSLFLALMHGSLGVCVOMELIYVER-QARLI 401
DB 433 IEVIFRRSSVSFVFSNKRILAHNLPFLILLFSTMGFYFIIRPLFYVFPVMSKLENGKFL 492
QY 402 QE--SPTLSKLA-----ITVLOPFYLL-----VOQTHMLEMGVSMAPCLFTDKML 448
DB 493 YELISSFIFLGLFIYHVIWIKOPFFRFLKFKIKIILH-----KFLINGW---- 539
QY 449 KYVKSIFYLGHIFPLSLFLPYI 472
DB 540 -----YF---DFEYKILFLTHPYL 554

```

RESULT 2

```

NUOL_BUCAI
ID NUOL_BUCAI STANDARD; PRT; 614 AA.
AC P57262;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE NADH-quinone oxidoreductase chain L (EC 1.6.99.5) (NADH dehydrogenase
  I, chain L) (NDH-1, chain L).
GN NUOL OR BU164.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
  symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
  Enterobacteriaceae; Buchnera.

```

```

OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RA MEDLINE=20445173; PubMed=10993077;
RX Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
  Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -I- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
  sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
  electrons transferred, four hydrogen ions are translocated across
  the cytoplasmic membrane), and thus conserves the redox energy in
  a proton gradient (by similarity).
CC -I- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -I- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNITS NUOA, H, J,
  K, L, M, N CONSTITUTE THE MEMBRANE SECTOR OF THE COMPLEX (BY
  SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -I- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
  OF CHLOROPLASTS OR MITOCHONDRIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AP001118; BAB12882.1; -
CC DR InterPro: IPR003916; NADhub_oxred5.
CC DR InterPro: IPR001750; Oxidored_q1.
CC DR InterPro: IPR001516; Oxidored_q1_N.
CC DR Pfam: PF00361; oxidored_q1_N.
CC DR Pfam: PF00662; oxidored_q1_N; 1.
CC DR PRINTS: PR01434; NADHDHGNASE5.
CC KW Oxidoreductase; NAD; Quinone; Transmembrane; Complete proteome.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 33 53 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 271 291 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 372 392 POTENTIAL.
FT TRANSMEM 410 430 POTENTIAL.
FT TRANSMEM 455 475 POTENTIAL.
FT TRANSMEM 492 512 POTENTIAL.
FT TRANSMEM 533 553 POTENTIAL.
FT TRANSMEM 593 613 POTENTIAL.
SQ SEQUENCE 614 AA; 70881 MW; 18CCCD2DFCAFE2780 CRC64;

```

Query Match 4.4%; Score 114; DB 1; Length 614;

Best Local Similarity 20.9%; Pred. No. 0.44; Indels 124; Gaps 20;

Matches 94; Conservative 66; Mismatches 165; Indels 124; Gaps 20;

```

QY 26 LSLNKLATSLGASEQAL---RLIISIFLGYPPALFYRHYLFYKETYLIHLPHFTGSLGIA 82
DB 68 LSLNEFKIDPGEFPLDGLSLMFLVITGGLIHIFSSMYMKKEQ--SRFAAYTNLPLFA 125
QY 83 YNFGNQLYHSLCTIVLOFLIRLMGRTTAVLTTCFQMAVLLAGYYTATGNYDIKMT 142
DB 126 SW-----SVVLADNLFEMYLGMGEVSV-----CYLLIGFYTELKNN--LCA 167
QY 143 MHCVLT-----LKLIGLVAVDPDGKDNSSLEQOKYAIRGVSPLLEVAGSFYFAGF 197
DB 168 FKAFLITVSVDFLIMGFILY---REFSNFNEQEIF---LSSFLNVNPFYLL----- 215
QY 198 LVGQPSNMHWMLVQGLIDIPKIPNSIIPALKRLSLGLFYLVGYTLLSPHITEDYL 257

```

Db 216 -----DITLFL--LIGVIGK--SAQLPLOTWSDA---WVGPPYSALHHAATMV 259
 QY 258 TEDYDNHFWERCMYLMIGKFLVYKYVTCMLTEGVCLTLGTFNGFEERKAKWDACA 317
 Db 260 TA-----GYLLARTHFLFLTPGLIYVGL----- 285
 QY 318 NMKWLFEETNREFGTGASFNININAMVARYIFRKLFLGKELSGSLFLFL-----W 373
 Db 286 -----IGTLTLIVSSISALVOKDI--KRI--LAVSTMSQ--IGVWFLALGVKAW 328
 QY 374 HGLHSGLVCFQWFLVIYVERQAARLIQESPTLSKLAITVLOPPLYVVOQTTHLMFLMG 433
 Db 329 SAAITHLIMHAIFFALLFLSAGSLIKSCKEKNIFKMGKRLKODPLFX-----ISFVVG 383
 QY 434 YSMTAFCLFTWDMKMLKYKYSIYFLGHIF 462
 Db 384 ASLVSPFLIT-----AGFYSKNLI 404

RESULT 3
 YA71_SCHPO STANDARD; PRT; 588 AA.
 AC 009758; Q9C025; (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 41, Last annotation update)
 DE Hypothetical protein C24H6.01c in chromosome 1.
 GN SPAC24H6.01C OR SPAB21F2.01
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Squares J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentile S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Mablett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voickart G., Aert R., Robben J., Grymopiriz B.,
 RA Wellens I., Vanstrele E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesl D., Hilbert H.,
 RA Borzyl K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu A., Dreano S., Gloux S., Lelaure V., Motier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt K., Moore K., Hurst S.M.,
 RA Lucas M., Rocher M., Gallard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.D.,
 RA Cerutti L., Lowe T., McComble W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO YEAST YGL084C.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: Z54142; CA90845.2; --
 DR EMBL: AL590562; CAC36890.1; --
 DR GenBank: SPAC24H6.01c; --
 DR Pfam: PF03062; MBOAT1; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 53 73 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 FT TRANSMEM 192 212 POTENTIAL.
 FT TRANSMEM 301 321 POTENTIAL.
 FT TRANSMEM 344 364 POTENTIAL.
 FT TRANSMEM 374 394 POTENTIAL.
 FT TRANSMEM 445 465 POTENTIAL.
 FT TRANSMEM 466 486 POTENTIAL.
 FT TRANSMEM 513 533 POTENTIAL.
 FT DOMAIN 22 40 SER-RICH.
 SQ SEQUENCE 588 AA; 69089 MW; 757AB685494B8B7A CRC64;

Query Match 4.2%; Score 109; DB 1; Length 588;
 Best Local Similarity 16.9%; Pred. No. 0.98; Mismatches 198; Indels 122; Gaps 18;
 Matches 83; Conservative 89; Mismatches 198; Indels 122; Gaps 18;
 QY 37 ASBQALILISIFLGYPPALFYHYLYFKET---YLHLEHTFTGSLIAYFNGQLYH 92
 Db 114 APENNNITLLVIVVFLRLVRLVAVPTNTNDELAIKNVRLCFSLLFLVYGVGVY 173
 QY 93 SLDCIVQFILLKMGTTTAVLTTFQFQNAVYLLAGYYT-----ATGN 136
 Db 174 VLTIALINYLISLKLKSNINPLTLTWLDSVVFKEYFAYCKRSSLHPGLGFLDQYTG 233
 QY 137 YDKKTPHCVLTKLIGLVNDY-----PDGQDNLS--SEOKVYIRG 180
 Db 234 LE-RWVYLENTMLRLSEFMDYWSLKHSEKNTLIFDKDEPTTLTRERVDYCLD 292
 QY 181 VP-SLELVAGSYFYGAFLVGPQFSNMHMYKLVOGELIDIPGKIPNSIIPALKSLGLF 239
 Db 293 EDVNLKNFLYIYFAPLYLGLISFNFMFSQ-----KYPT--VSLTKYRNL--- 338
 QY 240 YLVGYTLLSHITBEDYLLTDDY-----DNHPFWRCMYLMIGKFLV-Y 282
 Db 339 -LVAIRFLVCVLTMEFLHAYVYTAISKDNMNOYSAVESAMISFVLPW-TWLKLI 396
 QY 283 KYVTCMLTEGV-----CITGLGFGFEERKAKWDACAANMKWLPETNRFGT 334
 Db 397 RLFLNLSLIDIDEPENIVRCMCNNTSAVG-----WRA----- 430
 QY 335 ASFNININAMVARYIFRKLFLGKELSGSLFLFLMGLH-----SGLVCFQWFL 389
 Db 431 --WHRSENRRLIRIYVPLGSGNSHSLINLFIITFVALMWDISLWELFAKMWLI-----VL 483
 QY 390 IYVERQAARL-----IQESPTLSKLAITVLOPPLYVVOQTTHLMFLMGYSMTAFCLFTW 444
 Db 484 FILPERLCFMSRRTGTLKPYRYISGFAALINYMICNLIGFVAGIDIGKVLVVSF 543
 QY 445 DKMLKYKYSIYF 456
 Db 544 FLTLKGRKFLYF 555
 RESULT 4
 CYB_GARNE STANDARD; PRT; 360 AA.
 ID CYB_GARNE
 AC 079203;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome b
 GN MTCYB OR COB OR CYTB.
 OS Gasteria neriis (Grey-backed storm petrel).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Archaeosauria; Aves; Neognathae; Procellariiformes; Procellariidae;
OC Hydrobactinae; Garrodia.
OX NCBI_TaxID=79649;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate GBS-1;
RX MEDLINE=99003708; PubMed=9787440;
RA Nunn G.B., Stanley S.E.;
RT "Body size effects and rates of cytochrome-b evolution in tube-nosed seabirds";
RL Mol. Biol. Evol. 15:1360-1371(1998).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: AF076056; AAC68613.1;
CC InterPro: IPR005798; Cytochrome b6_C.
CC InterPro: IPR005797; Cytochrome b6_N.
CC Pfam: PF00032; Cytochrome b_C; 1.
CC Pfam: PF00033; Cytochrome b_N; 1.
CC PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
CC PROSITE: PS00193; CYTOCHROME_B_CO; 1.
CC KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane; Heme.
CC METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
CC METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
CC METAL 183 183 IRON 1 (HEME B562 AXIAL LIGAND).
CC METAL 197 197 IRON 2 (HEME B566 AXIAL LIGAND).
CC FT METAL
CC SQ SEQUENCE 380 AA; 42686 MW; 929F609033BC58F CRC64;

Query Match 4.1%; Score 107; DB 1; Length 380;
Best Local Similarity 20.8%; Pred. No. 0.87;
Matches 75; Conservative 47; Mismatches 106; Indels 132; Gaps 21;

OY 199 VGPOFSNMH-YMKLVQGLIDIPKIPNSI-----IPALKRLSLGLFVLGVYTLSPHIT 252
DB 1 MAPNLRKSHPLKIMNLSLIDLP--TPSNISAMWNFGSLGLCLVTQILTLG-LLLATHTY 57
OY 253 EDYLLTEDYDNH-----PFWRFCMYM-----LIWGEVLYKYVTCW 288
DB 58 ADTLTAFSSVHTCRNVQYGLINLNLHANGASFFPITCIYHIGRELYYGS---YLYKXTW 114
OY 289 LVTBVCILTLGL--GFNGFEKKGAKWDACANKWV-----LFTNPRFTGTIASFN 338
DB 115 -NTGILLTLTLMATAFVGY---VLPW---GQMSFWGATVITNLSAIPYIGQITIVW- 164
OY 339 INTNANVARYIFKRLKFLGNKELSGLSLFLALHGHGSLVGFQMEPLIVIERGAA 398
DB 165 -AMGFSV-----DNPLITRFALHFL-----LPEMAGLTLIHL 198
OY 399 RLIDESPILSKLAATVLAQ--PF--YLVQOTIHLFNGYSNTAPCLFTWD----- 445
DB 199 TFLHSSGNNPLGLVSNCDKIPFHYFSLKDTLGMFLPLITLTLALSPNLGDPENFT 258
OY 446 -----KMLKYKSYIF-----LGHIFPLS-----LFIPLYIKA 475
DB 259 PANPLVTPPHIKPEW-----YFLFAVAILRSIPNKLGVLAALASVLIIFLSPILHKS 311

```

```

RESULT 5
CYB_PELGA STANDARD; PRT; 380 AA.
ID CYB_PELGA
AC 079218;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cytochrome b
GN MTCYB OR COB OR CYTB.
OS Pelecanoides garnoti (peruvian diving petrel).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archaeosauria; Aves; Neognathae; Procellariiformes; Pelecanoididae;
OC Pelecanoides.
OX NCBI_TaxID=79637;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate PDP-1;
RX MEDLINE=99003708; PubMed=9787440;
RA Nunn G.B., Stanley S.E.;
RT "Body size effects and rates of cytochrome-b evolution in tube-nosed seabirds";
RL Mol. Biol. Evol. 15:1360-1371(1998).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: AF076073; AAC68630.1;
CC InterPro: IPR005798; Cytochrome b6_C.
CC InterPro: IPR005797; Cytochrome b6_N.
CC Pfam: PF00032; Cytochrome b_C; 1.
CC Pfam: PF00033; Cytochrome b_N; 1.
CC PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
CC PROSITE: PS00193; CYTOCHROME_B_CO; 1.
CC KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane; Heme.
CC METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
CC METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
CC METAL 183 183 IRON 1 (HEME B562 AXIAL LIGAND).
CC METAL 197 197 IRON 2 (HEME B566 AXIAL LIGAND).
CC FT METAL
CC SQ SEQUENCE 380 AA; 42653 MW; 7836F01B6A17840 CRC64;

Query Match 4.1%; Score 107; DB 1; Length 380;
Best Local Similarity 20.5%; Pred. No. 0.87;
Matches 75; Conservative 47; Mismatches 100; Indels 144; Gaps 21;

OY 199 VGPOFSNMH-YMKLVQGLIDIPKIPNSI-----IPALKRLSLGLFVLGVYTL 246
DB 1 MAPNLRKSHPLKIMNLSLIDLP--VPSNISAMWNFGSLTLITLITLTLGL-----L 51
OY 247 LSPHITEDYLLTEDYDNH-----PFWRFCMYM-----LIWGEVLY 282
DB 52 LAMHYTADDTLAFSSVAHTCRNVQYGLINLNLHANGASFFPITCIYHIGRELYYGS---Y 108
OY 283 KYVTCWLVTEGVCILTLGL--GFNGFEKKGAKWDACANKWV-----LFTNPRFTG 332

```

DB 109 LYKWT--NTGVILLTLMATAFVGY-----VLPW---GQMSFGATVITNFSAPYIGQ 159
 QY 333 TIASNTNTMNVARYIFRKLKLNKELSGSLFLFLMHGSLGVLVCQMELVI 392
 DB 160 TLVEW-----AMGGSV-----DNPTLTFPALHFL-----LPEMTVG 192
 QY 393 VERQARLIQESPTLSKLAITVLO--PF--YVLVQOTIHLFMGYSMTAFLPTWD--- 445
 DB 193 LSMHILTLHSGSNPLGIVSDCKIPHPYFSIKDLIGFTLMLPLTTLALFSPNLLG 252
 QY 446 -----KWLKVKYSIYF-----LGHIFPLS-----LPLFL 469
 DB 253 DPENTPANPLITPPHKEW-----YLFPAVALNSIPKLGVALASVLILFLA 305
 QY 470 PYIHK 475
 DB 306 PFLHKA 311

RESULT 6

HYBB_ECOLI
 ID HYBB_ECOLI STANDARD; PRT; 392 AA.
 AC P37180;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable Ni/Fe-hydrogenase 2 B-type cytochrome subunit.
 GN HYBB OR B2995.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / TGI;
 RX MEDLINE=94292472; PubMed=8021226;
 RA Menon N.K., Chateaus C.Y., Dervartanian M., Wendt J.C.,
 RA Shanmugam K.T., Peck H.D. Jr., Przybyla A.E.,
 RT "Cloning, sequencing, and mutational analysis of the hyb operon
 RL encoding Escherichia coli hydrogenase 2.";
 RL J. Bacteriol. 176:4416-4423(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blactner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- (FUNCTION: PROBABLE B-TYPE CYTOCHROME.
 CC -1- (SUBCELLULAR LOCATION: Integral membrane protein.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use. By non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U09177; AA021590.1;
 DR EMBL; U28377; AA69162.1; ALT_SEQ.
 DR EMBL; AE000382; AAC76031.1;
 DR PIR; A65086; A65086.
 DR Ecocyc; E01800; hydB.
 DR InterPro; IPR005614; NtFD.
 DR Pfam; PF03916; NtFD; 1.
 DR Transmembrane; Electron transport; Heme; Complete proteome.
 FT TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 35 55 POTENTIAL.

FT TRANSMEM 59 79 POTENTIAL.
 FT TRANSMEM 91 111 POTENTIAL.
 FT TRANSMEM 134 154 POTENTIAL.
 FT TRANSMEM 169 189 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 250 270 POTENTIAL.
 FT TRANSMEM 282 302 POTENTIAL.
 FT TRANSMEM 334 354 POTENTIAL.
 FT TRANSMEM 356 376 POTENTIAL.
 FT CONFLICT 306 307 NV -> KL (IN REF. 2).
 SQ SEQUENCE 392 AA; 43574 MW; CECF83746BCB8BFO CRC64;

Query Match 4.1%; Score 107; DB 1; Length 392;
 Best Local Similarity 21.0%; Pred. No. 0.89;
 Matches 89; Conservative 50; Mismatches 135; Indels 150; Gaps 21.

QY 79 LSIAYENFGNQLYHSLCTIYQPLILRMGTTITAVLTTCFQNAVLYAGYYTA-TGNY 137
 DB 72 LAMAVYVENRGQYHPL-----VRPALLASLFG-----YSIGGLSITIDVGRY 113
 QY 138 DIKTMHPICVLTILGLAVDYPPDGKQNSLSEQQ-----KVAIRGPSLLEAVG 189
 DB 114 ---WNLPPF-----YIPGHFNVNSVLEPFAVCMITYGWMLBEPALPFLRG 158
 QY 190 FSYFGAVLVGPOFSMNNHYMLVOGELIDIPGKIPIINSIIPALKRLSLGLFYLVGYTLSP 249
 DB 159 WKV-----SLQRLNKMFPIALGALLP 181
 QY 250 HITF---DYLLTFDYDHPFW-----FRCMYMLIGKFLVKYVTCMLVTEGCIL 297
 DB 182 TMHOSMSGLMISAGYKVPIMQSYEMLPFLSLTLAPFMGPSI-----VIFEGSLVQ 233
 QY 298 TGLGFNGFEKQKAKMDACAMKMWLPETNRFCTISFNINNAMVARYIFRKLKPLG 357
 DB 234 AGLKNGGPDERS-----Lr---VKLTNTFS-----VLLAIFVLRP-G 267
 QY 358 NKEISQGLSLFLALMHGSLGVLVCQMEFLIVVERQARL--IQSPFLSKLAITV 415
 DB 268 ELIYRDLKSLAFAG-----DFYSVMFIEVLMLFLVLRVANVRDMSMLPLSALSA 321
 QY 416 LQPFYLVQOTIHLFMGYSMTAF-----CLF-TWDRMLKV--YKSIYFGLHIFFLSL 466
 DB 322 L-----LGCATWR-LTYSLVAFNPGGYAVFPTWELLISIGVAIEICAYVILRL 373
 QY 467 FILP 470
 DB 374 PILP 377

RESULT 7

YCURB_SCHPO
 ID YCURB_SCHPO STANDARD; PRT; 537 AA.
 AC 059831;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative amino-acid permease C965.11C.
 GN SPOCC965.11C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsle K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittsch E.,
 RA Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,
 RA Skelton J., Simmons M., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Skelton J., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodard J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas R., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revelata J.L., Moreno S., Armstrong J., Forstberg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: Belongs to the amino acid permease family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL023590; CAA19071.1; -
 CC PIR: T41664; T41664
 DR GeneDB: SPombe; SPCC965.11c; -
 DR InterPro: IPR002293; AA/rel_permease1.
 DR InterPro: IPR004840; AAC_permease.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; FALSE_NEG.
 DR Hypothetical protein; Transport; Amino-acid transport; Transmembrane.
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 157 177 POTENTIAL.
 FT TRANSMEM 183 203 POTENTIAL.
 FT TRANSMEM 219 239 POTENTIAL.
 FT TRANSMEM 267 287 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 393 413 POTENTIAL.
 FT TRANSMEM 462 482 POTENTIAL.
 FT TRANSMEM 537 AA; 59431 MW; C6022537E9AE1CF CRC64;
 SO SEQUENCE

Query Match 4.1%; Score 106.5; DB 1; Length 537;
 Best Local Similarity 20.1%; Pred. No. 1.4;
 Matches 67; Conservative 58; Mismatches 118; Indels 91; Gaps 17;

QY 5 AAGDECTVAL--AGVLGGFDELSTNKATLSGASEQRLRLISFLGPPALFRNHL 62
 DB 222 AAGFGGIVAFYVAGVFSGVESVMT--AAESKNPKKAIPLAV----- 263
 QY 63 FKEETLHLHFTFTGLSIAY-----FNFQNLVSLLCIYGLILRLMRTIT 112
 DB 264 --RQTFW-RILVYFISISISYGTIVANDPNLSSGSKTLSPMTIIMAGNHNAGDPFN 320
 QY 113 AVLTFFCFQMAVLLAGYVYATGNVD-----IKWTMPCVLLTKLIG 154
 DB 321 AVLTTC--LSSINSIGYISRLVYMLAKDGMAPKIFKRVKRGVPMVAHVHLEGFSL 378
 QY 155 LAVVDYDGGKDNSSLSSEQKAIKRVPSLLEVAGSYFGALVLPQSSNHHY--K 210
 DB 379 I-MNVSSTG-----AVKAYGYIINLAGS-----AFIVTATIIFFVFRFRRW 419
 QY 211 LVQGG-ELIDIPKIPNSIIPALKRLSLGLFYLV--GYTLLSPHITEDYLLTEDYDNHPW 267

DB 420 VKQYALSDLPFKSPDYLPFQILGIVGIIITLVQGWTFVFPFAGDPV--DAYLLBLF 477
 QY 268 FRCMMLWKGFLVLYKYV--TCMLVTEGVCILLTG 299
 DB 478 F-----VTM-----LSYKPIKKTWVSYEDMDEING 503

RESULT 8
 COX1_LOCM1 STANDARD; PRT; 512 AA.
 ID COX1_LOCM1
 AC Q36421;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1).
 GN COL.
 OS Locusta migratoria (Migratory locust).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.
 OK NCBI_TaxID=7004;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96139026; PubMed=8587138;
 RA Flock P.K., Rowell C.H.F., Gellisen G.;
 RT "The sequence, organization, and evolution of the Locusta migratoria
 RT mitochondrial genome.";
 RL J. Mol. Evol. 41:928-941(1995).
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 CC AND COPPER B.
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -1- PATHWAY: Respiratory chain; terminal step.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X80245; CAA56527.1; -
 DR PIR: T11467; T11467.
 DR HSSP: P00396; 20CC.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASE1.
 DR PROSITE: PS00077; COX1; 1.
 DR Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;
 KW Respiratory chain; Inner membrane.
 KM METAL 60 60
 FT METAL 239 239 IRON (HEME A AXIAL LIGAND) (PROBABLE).
 FT METAL 243 243 COPPER B (PROBABLE).
 FT METAL 289 289 COPPER B (PROBABLE).
 FT METAL 290 290 COPPER B (PROBABLE).
 FT METAL 375 375 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
 FT METAL 377 377 IRON (HEME A AXIAL LIGAND) (PROBABLE).
 FT METAL 377 377 IRON (HEME A AXIAL LIGAND) (PROBABLE).
 FT CROSSLINK 239 243 1'-histidyl-3'-tyrosine (By similarity).
 SO SEQUENCE 512 AA; 56814 MW; 98031A229B00511C CRC64;

Query Match 4.1%; Score 106; DB 1; Length 512;
 Best Local Similarity 21.7%; Pred. No. 1.4;

Matches 102; Conservative 55; Mismatches 155; Indels 158; Gaps 25;

QY 88 NOLYSLICIVLQFILRLMGRITAVLTTCFCOMAYLAGYATATGVYDK----- 140

DB 50 DQLVVIT-----TAAAFWIFPMWPMIGG-----GNLVPLMIGAD 90

QY 141 -----WTMPHCVTTLKLGAVDYFDGKDONSLSSEQOKYAIRGPSL----- 184

DB 91 MAFPMNNMSFWLLBPSS-LTLLMSSVVD-----NAGAGMTVY-----PLASVIA 136

QY 185 -----LEVAGPSYFYGA--FLVGPQPSNMHYKLVQGLIDIPGKIPNSIIPA 230

DB 137 HSGASVDLAIFSLHAGVSSILGAINFTTAINMSSNMNTLDQTELF-----VMSVATTA 191

QY 231 LKRLSLGLFYLVG-YTLSPHITEDYLITEDYD-----NHPWF-----RCMYMLIM 276

DB 192 L-LLLSLPYLAGATLML-----LTDRLNLTSSFPDPAGGDPILYQILFMFGHPEYIIL 247

QY 277 GKEVLVYKVTW-----LVTEG-VCIITGLFNGFEEKKAKWADACANMKVWL----- 323

DB 248 PGFGIISHIVCOESKIESFGTIGMIVAMLSIGLMGF-----IVMAHMT 293

QY 324 -----PETMPRTGTIASFRININ-----AMVAVYIKRLKFLGNKELSGSLFLALMHG 375

DB 294 VGMDVDTAAYFTSATMIIAIVPGIKVFSWMAATLYGTKEF--NPPLMALGFIFLFTMG 351

QY 376 LHSGLVCOEFLIVIVERQARLIQSSPTLSKLAATVLOPFYVVOQTTHLMFGYS 435

DB 352 LTG--LVLANSLDVLHD-----TYVVAHFYVLSMGAVFAIMGSI 392

QY 436 MTAFLCLFTW---DKMLKYKSIYFLGHIFLSLFI-LPYTHKAMVPRK 480

DB 393 IOWYLFPLGLTWNKMKLKIQFTIMFIG---VNLFFPHQHFLGLAGMRR 438

RESULT 9

CYB_PTEHY STANDARD; PRT; 380 AA.

AC 079224;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Cytochrome b.

GN MTCYB OR COB OR CYTB.

OS Pterodroma hypoleuca (Bonin petrel).

OC Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauaria; Aves; Neognathae; Procellariiformes; Procellariidae;

OC Procellariinae; Pterodroma.

OX NCBI_TaxID=79640;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Isolate Bonin-1;

RX MEDLINE=99003708; PubMed=9787440;

RA Num G.B., Stanley S.E.;

RT "Body size effects and rates of cytochrome-b evolution in tube-nosed seabirds".

RL Mol. Biol. Evol. 15:1360-1371(1998).

CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).

CC -1- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 (By similarity).

CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).

CC -1- SIMILARITY: Belongs to the cytochrome b family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC -----

DR EMBL: AF076079; AAC68636.1; -

DR InterPro: IPR005798; Cytochrome b6.

DR InterPro: IPR005797; Cytochrome b6 N.

DR Pfam: PF000032; Cytochrome b c1.

DR Pfam: PF000033; Cytochrome b N1.

DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.

DR PROSITE: PS00193; CYTOCHROME_B_OO; 1.

DR Electon transport; Mitochondrion; Respiratory chain; Transmembrane; Heme.

KW METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).

FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).

FT METAL 183 183 IRON 1 (HEME B566 AXIAL LIGAND).

FT METAL 197 197 IRON 2 (HEME B566 AXIAL LIGAND).

SQ SEQUENCE 380 AA; 42668 MW; A8217408D11FB49B CRC64;

Query Match 4.1%; Score 105; DB 1; Length 380;

Best Local Similarity 21.0%; Pred. No. 1.2; Indels 126; Gaps 20;

Matches 75; Conservative 46; Mismatches 110;

QY 199 VGPOFSNMNH-YMKLVQGLIDIPGKIPNSI-----IPALKRLSLGLFYLVGTYTLSPHIT 252

DB 1 MAPNLRKSHPLRLKVNNSLIDLP--PNTISAMNPFSLGICLMTQTLTG-LLLAMHYT 57

QY 253 EDYLTEDYDNH-----PMPFRCKYMLIWGKFLV--KYVTCWLVT 291

DB 58 ADTLTAFSSVAHTRGVQYGLIRNLHANGASFFPICYLHIGRGFYGSYLKETE--N 115

QY 292 EGVCLITGL---GNGFEEKKAKWADACANMKW-----LFENPRTGTIASFRININ 341

DB 116 TGVMLLTLMATAVGV---VLPM--GQMSFWGATVITNLFNAIPYIGQTLVEW----- 164

QY 342 NAMVAVYIFKRLKFLGNKELSGSLFLALMHGSHGLVCOEFLIVEROARLI 401

DB 165 -AMGFSV-----DNPTLTRPFAHLT-----LPFYAGTLIHILTL 201

QY 402 QESPTLSKLAATVLO--PY--YVLVQQTTHLMFGYSMTAFCLFTWD----- 445

DB 202 HESGSNNPLGIIVSNCDKIPFHPYFTLKIDLFEMFLPLTLTALFSPPLDGPENFTPAN 261

QY 446 -----KMLKYKSIYF-----LGHIFPLS---LFIPIYTHKA 475

DB 262 PLITPPIKPEW-----YPLFAVAILRSIPNKLGGVLAASVLIILFVPLHKA 311

RESULT 10

COX1_APILI STANDARD; PRT; 521 AA.

AC P20374;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1).

GN COI.

OSApis mellifera ligustica (Common honeybee).

OC Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyigota;

OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;

OC Apidae; Apis.

OX NCBI_TaxID=7469;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Thorax;

RX MEDLINE=90136028; PubMed=2559293;

RA Crozier R.H., Crozier Y.C., Mackinlay A.G.;

RT "The CO-I and CO-II region of honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates".

RL Mol. Biol. Evol. 6:399-411(1989).

RN [2]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Thorax;
RX MEDLINE=93114603; PubMed=8417993;
RA Crozier R.H., Crozier Y.C.;
RT "The mitochondrial genome of the honeybee Apis mellifera: complete
RL sequence and genome organization."
CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
CC AND HEME A OF SUBUNIT 1 TO THE BINETALLIC CENTER FORMED BY HEME A3
CC AND COPPER B.
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23409; AAA18476.1; -
DR EMBL: L06178; AAB96799.1; -
DR PIR: A32431; A32431.
DR HSSP: P18401; 1FTF.
DR InterPro: IPR000883; COX1.
DR Pfam: PF00115; COX1.
DR PRINTS: PR01165; CYCOXIDASE1.
DR PROSITE: PS00077; COX1.
KW Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;
KW Respiratory chain; Inner membrane.
FT METAL 238 59 IRON (HEME A AXIAL LIGAND) (PROBABLE).
FT METAL 242 242 COPPER B (PROBABLE).
FT METAL 288 288 COPPER B (PROBABLE).
FT METAL 289 289 COPPER B (PROBABLE).
FT METAL 374 374 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
FT METAL 376 376 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
FT CROSSLINK 238 242 1'-histidyl-3'-tyrosine (by similarity).
SQ SEQUENCE 521 AA; 59293 MW; 2149417AC981CE64 CRC64;

Query Match
Best Local Similarity 4.1%; Score 105; DB 1; Length 521;
Matches 93; Conservative 61; Mismatches 148; Indels 178; Gaps 24;

QY 88 NOLYHSLCLIVLQFLRLMGRITVAVLTFFCQMAVYLAGYVYRTGNYDIK----- 140
DB 49 DOIYNTIV-----TSHAFLMIFFWVMPFLIGF-----GWLPLPLMGSPD 89
QY 141 -----WTMPCVLTLLKLGAVDYPDGKQNSLSSEQKAYIRGVPSL----- 184
DB 90 MAFPRMNNISFWLLPFSILMELLSLN-----FYSPGCGWTVYPPLSAYLYHSBDSVDFALF 146
QY 185 -LEVAGFSYFYGA-----FLVGPQPSMNHMYMKLVGSELIDIPGKIPNSIIPALKSLSG 237
DB 147 SLHMSGISIMGSLNLMVTIMMKNSMY-----DQSLSPWVFTAILILNS 196
QY 238 LYLIVG-YTLSPHTEDYLTEDVD-----NHPWF--RCMTMLWGKRVLYK 283
DB 197 LPLVLAITML-----LFDNRNFTSFDPMGSDPILYOHFLWFGHBEVYVILILPGGLIS 253
QY 284 VYTCMLVTEG-----VCILTGFGNGFEEKGAKMDCAMKMWL----- 323
DB 254 H1--VMNESGKKEITGNLSMTIAMLGIGLGF-----IWAHMTFTVGLD 296

```

```

QY 324 FETNREFTGTIASFNINTN-----AWVARYIFKRLKELGNKLSOGLSLFLALWHGLHSG 379
DB 297 VTRAVYFISATMIIVPTGIKIVFSWLATYHSGSKL-----LNISILW--SLG 341
QY 380 YLVCGMEFLVIVRQARLRQESPTLSKLAIRVLOPFY-----LV 423
DB 342 FIMLFTIGTLGIM-----LSNSIDILIDTYVGVGHFHYVLSMGAVFAL 388
QY 424 QOTIHW--LFGYSWTFACFLTWDKLVKYSIYFLGHIFSLFLIPYIKHM-VPRK 480
DB 389 SSFIHWYELITGLNLNT-----KWLKIQIFMWTIG--VNLTFQGHFLGLMSMRK 437

RESULT 11
NUSM_ALBICO STANDARD; PRT; 545 AA.
ID NUSM_ALBICO
AC P48918;
DT 01-FEB-1996 (rel. 33, Created)
DT 01-FEB-1996 (rel. 33, Last sequence update)
DT 01-FEB-1996 (rel. 33, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5
OS Albinaria coerulea (land snail).
OC Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigurethra; Clausiliidae; Clausiliidae; Altophinae; Albinaria.
OX NCB1_Taxid=42349;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96120351; PubMed=7498775;
RA Hatzoglou E., Rodakis G.C., Iecanidou R.;
RT "Complete sequence and gene organization of the mitochondrial genome
RT of the land snail Albinaria coerulea."
RL Genetics 140:1353-1366(1995).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X83390; CAA58296.1; -
DR PIR: SS9143; SS9143.
DR InterPro: IPR003916; NADHUB_OXRED5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; oxidored_q1.1.
DR Pfam: PF00662; oxidored_q1.1.
DR PRINTS: PR01434; NADHUBGNASE5.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 545 AA; 61335 MW; B8801FDECA18563 CRC64;

Query Match
Best Local Similarity 4.1%; Score 105; DB 1; Length 545;
Matches 81; Conservative 42; Mismatches 126; Indels 86; Gaps 19;

QY 4 SAEDEGVVALAGVLSGFQELSLNKLATSGASEQALRLITGFLGYP-AL--FYRH 60
DB 231 SLMGVGSITCLLGSVALF-EYDKKV-IALSTLSQGVWVYSLSLNPYALALHLYGH 288
QY 61 YLFYEKTLIHNFHTFGSLAYFNGVQ-----LYHSLLCT-VLQFLRLMGRIT 111
DB 289 AMFKAMFLP-----GAGL-IIMNSYGTQDLRLGLSLYSSPTVSLTSLTSLWCLMGPF 341
QY 112 TA-----VLTFFCQMAVYLAGY-----YATGNYDIK----- 140
DB 342 VSSFSKHLLEKMLDNMNCNFTSMFVLGTLITGMYSIRLMKFLCKWGNNNKSPCYNMS 401
QY 141 WTMPCVLTLLK-----LIGLAVDYPDGKQNSLSSEQKAYIRGVPSLLEVAGFSYFYGA 196

```


Db 402 WOSKSMFPLAALAVSGLMSYLDSSYMTFSWSTNQNLIMGVLFISFGIWKFCN 461
 QY 197 F-----LVGP-QFSNMHYMK--LVQGEIIDIPIKIPISIPALKRLS-----LGTF- 239
 Db 462 FYPTLMSMMFLGPFPSYNLHTKSLIYMKRIDISISEPNVMGNLWSSWRWMSLFN 521
 QY 240 YLVGTTLSPHITEDYLTEDYDNHPFWRFCRYMLI 276
 Db 522 WLTNYMLVLT-----WFLLV-----WLMIMSLMW 545

RESULT 12

NOSM_RHIST

ID NOSM_RHIST STANDARD; PRT; 654 AA.

AC P50367;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 5 (BC 1.6.5.3).
 GN NDS OR NAD5.
 OS Rhizopus stolonifer (Rhizopus nigricans).
 OG Mitochondrion.
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 OC Rhizopus.
 NCBI_TaxID=4846;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DAOM 148428;
 RA Paquin B., Roewer I., Wang Z., Lang B.F.
 RT "A robust fungal phylogeny using the mitochondrially encoded nads
 protein sequence."
 RL Cam. J. Bot. 73:S180-S185(1995).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-stb.ch/announce/>
 or send an email to license@isb-stb.ch).
 CC
 CC EMBL; U17011; AAA9062.1; -.
 DR PIR; T14202; T14202.
 DR InterPro; IPR003916; NADHox oxidred.
 DR InterPro; IPR001750; Oxidored_q1.
 DR InterPro; IPR001516; Oxidored_q1_N.
 DR Pfam; PF00361; Oxidored_q1; 1.
 DR Pfam; PF00662; Oxidored_q1_N; 1.
 DR PRINTS; PR01434; NADHoxNASES.
 DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 654 AA; 72818 MW; 22D2E05B8D6E12D CRC64;

Query Match 4.1%; Score 105; DB 1; Length 654;
 Best Local Similarity 21.5%; Pred. No. 2.1;
 Matches 87; Conservative 59; Mismatches 164; Indels 94; Gaps 20;

QY 14 ALAAGVGGFQELSLINKATISGASEQALRLIISIFLGIPPLF-----YRHVLFYGE 66
 Db 289 ATTGLQNDLKRVIAYVSTCSQG-----LLFVCGLSQYNVALFHLVNHAMFKALLFFSA 343
 QY 67 TYLIHLFH-----TFTGLS-LAVNFGNQLYSLICYLOFLIRLMQRTTAVLT--- 116
 Db 344 GSVIHAMNDEODLRKFGGSLRLPFTYSMVVIGSLSMALPPL-IGFYSKDLITELAYGH 402
 QY 117 -TFCFOMAYLLAGYYTATGNDYDKMTWCHVLTLLIGLAVDYDGGK-DONSLSSEQQ 174
 Db 403 YPSGSLVYVWLASVAAVFTAMYSIR-----SLVLTFLG-----YENGPKINYNNTI--HE 449
 QY 175 KVAIGVPSLELVAGSIFYG-----AFVGPQFSNMHYMKLVQGEI 217
 Db 450 ADLMAIAP-LVLAVALSIFFGVYTKDLFVGMGTDFYNNALFTHPHSI-----LV 498

QY 218 DIPKIPNSI--IPALKSLGLFVYGVYLLSPHITEDYLLTE-----DYDNHPFWR 269
 Db 499 DTEFGLPMSMKRFLPLISGL-LGTFGVALIYWIPELPPKFIISTKRGRIYFFPNQKYFD 557
 QY 270 CMYMLINKFPLVKKYVTWLTVEGVCILITGFGNFEKGAQKADCAKMYWLEPTPR 329
 Db 558 NIYNMLNKKFLNFGYTTKILDRGAIELVG-----PYGLVNVFKSASNNVSGLDSC-- 608
 QY 330 FTGTIASFNINMNAVARIYFPRLLKFLGNKELSGQLSLPLALW 373
 Db 609 FIPTYAMTIFNGLLIFTLIF-----FIQDPR-----LFLULLM 642

RESULT 13

CYB_BRALA

ID CYB_BRALA STANDARD; PRT; 380 AA.

AC P92472; O79415;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome b.
 GN MTCYB OR COB OR CYTB.
 OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 NCBI_TaxID=7740;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96292550; PubMed=9628930;
 RA Spruyt N., Delarbre C., Gachealin G., Laudet V.
 RT "Complete sequence of the amphioxus (Branchiostoma lanceolatum)
 RT mitochondrial genome: relations to vertebrates."
 RL Nucleic Acids Res. 26:3279-3285(1998).
 RN [2]
 RP SEQUENCE OF 186-380 FROM N.A.
 RA Gachealin G.;
 RT Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 complex (complex III or cytochrome b-c1 complex), which is a
 respiratory chain that generates an electrochemical potential
 coupled to ATP synthesis (By similarity).
 CC -1- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
 CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
 CC or b566) is high-potential and absorbs at about 566 (By
 CC similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-stb.ch/announce/>
 or send an email to license@isb-stb.ch).
 CC
 CC EMBL; Y16474; CAA76246.1; -.
 DR EMBL; Y09849; CAA70979.1; -.
 DR PIR; A71390; A71390.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00032; Cytochrome_b_c1.
 DR Pfam; PF00033; Cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_QQ; 1.
 KW Election transport; Mitochondrion; Respiratory chain; Transmembrane;
 HM Heme.

FT METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 183 183 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 197 197 IRON 2 (HEME B566 AXIAL LIGAND).


```

Db 117 SFLGISVLTGKGVOLINEARDG-----LKEQ-----LP-LHRLLYFLPEPTISSG 162
Qy 201 POFSMNHMYKLVOG-----ELIDIPGKIPNSIIPALKSLGLFY--LVGYTLIS-- 248
Db 163 P---IDRYRRVVKBOQAKMTKEEYADL-----LYGIKIKIFIGFLYKPIGIAINTYF 212
Qy 249 ----PHITEDYLLTDEDYDNHDFRCMWMILMGKFLVLYKYTCMLVTEGVCILTGL----- 300
Db 213 IMNLPALITHNKILG-----NLLYMGYSMYLFFDFGAYTMFANVGYINGIKSPE 262
Qy 301 GFNFEEBKGRKAKMACANMKWLEPTNPRFGITIASF---NINTNANVARIYIKRLKFL 356
Db 263 NFN-----KPFISKNIKDFENRWMSLSFWRDYPVFRFVFW 299
Qy 357 GNKE-----LSQGLSLF--LALWGHSGVLYVCFQWMEFLIVYERQARLIQESP 405
Db 300 MTKKKWIKNNMAVSNIGYFLFLFMGMVHGAPQYI----- 335
Qy 406 TLSKLAATVLPYFYLVQOTIHWLFMGYSMTAFCLFTWDKMLKYK 452
Db 336 -----IYGLYHVAVLMTGYNPF--EKMKKKYK 359

RESULT 15
CYB TINMA STANDARD; PRT; 379 AA.
AC 003553;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cytochrome b.
GN MTCYB OR COB OR CYTB.
OS Titamus major (Great titmouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Tinamiformes; Tinamidae; Tinamus.
OC NCBI_TaxID=30468;
RN [1]
RP SEQUENCE FROM N. A.
RA Lee K., Feinstein J., Cracraft J.;
RT "Phylogenetic relationships of the ratite birds: resolving conflicts
RT between molecular and morphological data sets.";
RL (in) Mindell D.P. (eds.);
RL Avian molecular evolution and systematics, pp.1-1, Academic Press,
RL New York (1997).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
CC or b566) is high-potential and absorbs at about 566 (By
CC similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U76056; AAB6130.1;
CC InterPro: IPR005798; Cytb b6 C.
CC InterPro: IPR005797; Cytb b6 N.
CC Pfam: PF00003; cytochrome_b_C; 1.
CC Pfam: PF00003; cytochrome_b_N; 1.
CC PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
CC PROSITE: PS00193; CYTOCHROME_B_OO; 1.
CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW

```

```

KW Heme.
FT METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 197 197 IRON 2 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 379 AA; 42797 MW; 60B9D75CBEA430AA CRC64;

Query Match 4.0%; Score 104; DB 1; Length 379;
Best Local Similarity 21.1%; Pred. No. 1.4; Indels 80; Gaps 11;
Matches 63; Conservative 41; Mismatches 114;

Qy 46 ISIFLGYPALFYRHHYLFYKETYLIHLFHTGSLAVFNGQLYHSLCYVLPILR 105
Db 93 ICIVLHIGRGLYGYSYL-YKETWNIGIILLTLMATAFVG-----VLPMQOMS 140
Qy 106 LMGRTITAVLTTRCFQNAVLLAGIYYATGNYD-----KATMPHCUTLKLIG 154
Db 141 FWGAT--VITNLSAIPYIGQALVEWAWGFSVDNPTLTRFPALHFLPLIVLSLSIH 197
Qy 155 LAVDPDGGKQDQNSLSEQQKVAIRGVPSLLEVAGF-----SYFYGAFL----- 198
Db 198 LTFHESGSNNPLGIISQDKIPHPYFTTKDMGLFTLMPPLTLTAFFPNFLGDEBNF 257
Qy 199 -----VGPOFSMNHMYKLVOGELIDIPGKIPNSIIPALKRSLGLPYLVGYTLISP-- 249
Db 258 TPANPLITPRHIKPEWYFLFAVAILRSIPNKLGG--VIALTASVLIF-----LSPL 308
Qy 250 HITEDYLLTDEDYDNHDFRCMWMILMGKFLVLYKYTCMLVTEGVCILTGLGFNPFEE 307
Db 309 HASKMRSLT-----FRMSQLF-----WLIANLILITWIGSQPVED 346

```

Search completed: November 21, 2003, 14:15:12
 Job time : 19 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:13:48 ; Search time 21 Seconds

(without alignments)
981.209 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 2580
Sequence: 1 MASSAEGDEGVVALAGVLQ.....ILPYHKAMPKREKKME 487

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	113	4.4	484	US-09-634-238-216	Sequence 216, App
2	105	4.1	2132	US-08-669-656A-6	Sequence 6, Appl
3	102.5	4.0	412	US-09-134-001C-4885	Sequence 4885, Ap
4	99.5	3.9	443	US-09-328-352-7069	Sequence 7069, Ap
5	98.5	3.8	243	US-09-252-991A-29870	Sequence 29870, A
6	98.5	3.8	401	US-09-107-532A-3878	Sequence 3878, Ap
7	97.5	3.7	448	US-09-328-352-8050	Sequence 8050, Ap
8	96.5	3.7	551	US-09-252-991A-32219	Sequence 32219, A
9	95	3.7	551	US-09-328-352-7176	Sequence 7176, Ap
10	93	3.6	793	US-09-328-352-7092	Sequence 7092, Ap
11	93	3.6	506	US-09-066-047-7	Sequence 7, Appl
12	93	3.6	793	US-09-107-532A-5141	Sequence 5141, Ap
13	92	3.6	380	US-09-097-889A-25	Sequence 25, Appl
14	92	3.6	470	US-09-098-079-25	Sequence 25, Appl
15	92	3.6	380	US-09-328-352-5397	Sequence 5397, Ap
16	91	3.5	535	US-09-117-860-18	Sequence 18, Appl
17	91	3.5	535	US-09-117-860-24	Sequence 24, Appl
18	90.5	3.5	1302	US-08-232-537-2	Sequence 2, Appl
19	89.5	3.5	488	US-08-928-692-11	Sequence 11, Appl
20	89.5	3.5	488	US-09-339-972-11	Sequence 11, Appl
21	89.5	3.5	795	US-09-198-452A-314	Sequence 314, App
22	89.5	3.5	832	US-08-677-734A-12	Sequence 12, Appl
23	89.5	3.5	832	US-09-097-053-12	Sequence 12, Appl
24	89	3.4	361	US-09-134-001C-4045	Sequence 4045, Ap
25	89	3.4	409	US-09-071-035-316	Sequence 316, App
26	89	3.4	433	US-09-071-035-314	Sequence 314, App
27	89	3.4	522	US-09-165-042-3	Sequence 3, Appl

28	89	3.4	720	US-09-252-991A-19581	Sequence 19581, A
29	88.5	3.4	331	US-09-107-532A-5276	Sequence 5276, Ap
30	88.5	3.4	594	US-09-107-532A-6929	Sequence 6929, Ap
31	88.5	3.4	962	US-09-328-352-7942	Sequence 7942, Ap
32	88	3.4	360	US-08-597-236-11	Sequence 11, Appl
33	88	3.4	360	US-08-746-682A-11	Sequence 11, Appl
34	88	3.4	664	US-09-268-140-2	Sequence 2, Appl
35	87.5	3.4	448	US-08-811-897A-18	Sequence 18, Appl
36	87.5	3.4	448	US-08-855-213-18	Sequence 18, Appl
37	87.5	3.4	448	US-09-201-474-18	Sequence 18, Appl
38	87.5	3.4	467	US-08-811-897A-19	Sequence 19, Appl
39	87.5	3.4	467	US-08-855-213-19	Sequence 19, Appl
40	87.5	3.4	467	US-09-201-474-19	Sequence 19, Appl
41	87	3.4	474	US-09-790-838-4	Sequence 4, Appl
42	87	3.4	537	US-08-647-397-2	Sequence 2, Appl
43	87	3.4	834	US-08-677-734A-9	Sequence 9, Appl
44	87	3.4	834	US-08-677-734A-10	Sequence 10, Appl
45	87	3.4	834	US-09-097-053-9	Sequence 9, Appl

ALIGNMENTS

```
RESULT 1
US-09-634-238-216
Sequence 216, Application US/09634238
Patent No. 6544772
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
APPLICANT: Dekker, James
APPLICANT: Christensen, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
APPLICANT: Coolbear, Timothy
TITLE OF INVENTION: Polynucleotides, materials incorporating
FILE OF INVENTION: them and methods for using them.
FILE REFERENCE: 11000.104301
CURRENT APPLICATION NUMBER: US/09/634.238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422
SOFTWARE: FaSTSeq for Windows Version 4.0
SEQ ID NO 216
LENGTH: 484
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-09-634-238-216
Query Match: 4.4%; Score 113; DB 4; Length 484;
Best Local Similarity 20.4%; Pred. No. 0.0019;
Matches 100; Conservative 70; Mismatches 155; Indels 166; Gaps 23;
46 ISIFGYFALPRHYLYFK-ETVYIHLFHPFTGSLAYFNGNGLYSLICIVLQPL 104
85 LAYVAMTYWVHPIYLKQPOAILIALSWLFKG-NGPVTWVSMTVSLICLAFLL 143
105 RLMGRITL-----AVLTFFCFOM-----AYLLAGYYVYTAGVY-DIKMTPHCV 147
144 WLSSKGLTTLNRIGSMAGTAMFLMSILFIILAVTAPLWVKGHVATPRMGDIHTLP 203
148 L-----TLKILGLANDYFPGGQDQNSLSEQOKYAIRVP-SLLEVAG--FSYFGATLV 199
204 LNYFTTISMLVFVAV---GGAEKISPYVNNTKHKSKEPPLGLWLVLAGVAFALGSGM 259
200 GPQFSMNHYMLVQCELDIDPKIPNSI---PALKRLSLGLFVAGVTLSPHTBYL 256
260 GMLFNSHH-----LPTDLMANGPYTAPMLGKYIVG----- 291
257 LTEDYDHPFWRFCMYMLMGKPVLYKYVTCWLVTEGVCIITGLGFGNFEKGAKWAC 316
```

Db 292 -----GLFVI-----LYSIANALASISALAFS-----ID 315
QY 317 ANMKWLFETNPF-----TGTIASFNNTAMVARYFKRLKFL 356
Db 316 APKILSLADDPHFIRKRLRINKRKTPIINGYMTILVSLIITPAFGIGNNELYKWL 375
QY 357 GN-KELSGISSLFLALMHLGSLVLCFQMEFLIYVEROARLLIOESPPLSKLAATV 415
Db 376 LNLNSVVMPLRYLWFLAVYLLNQHLQEKQADYMLV-----KKRLGLMVG----- 421
QY 416 LQPFYLVVQOTIHLFMGYSMTAF-CLF-----TWDKRLKYKSIYFLGHIF 461
Db 422 -----RMCF-----FTAFACLLGMPLKRIYLANDPGTW--WFOV--GLNITTPVI 462
QY 462 FLSLFPLPYI 472
Db 463 LMLGLILPYI 473

RESULT 2
US-08-669-656A-6
; Sequence 6, Application US/08669656A
; Patent No. 6451554
; GENERAL INFORMATION:
; APPLICANT: Wood, John N.
; APPLICANT: Akopian, Armen N.
; TITLE OF INVENTION: Ion Channel
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike, P.O. Box 15437
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,656A
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Lisa D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-669-656A-6

Query Match 4.1%; Score 105; DB 4; Length 2132;
Best Local Similarity 19.4%; Pred. No. 0.13;
Matches 113; Conservative 74; Mismatches 194; Indels 200; Gaps 28;

QY 140 KWTMPHCVL-TKLIGLAVDPDGKDONSLSSBOOKAIRGVPSLLVAGFSYFYGAFL 198
Db 728 KNINIDCVIIVLSLELS-----ASKGSL-----VLKSLRLALDITGQKSLFSAQY 775
QY 199 VGPQFSNMHYKLVQ-----GELIDIPGKIPNSII-----PALKRLSLGLFY 240
Db 776 LNEPFRAGRAMSVSIVTSIELESKLCPCPLISFQAKYLIWECCKMRKFKMALPE 835
QY 241 LVGYTLSPHIT-----EDYLLTEBYDNHPFRMCMWLINCKFLYKXVTCW 288
Db 836 LVTDFPABLITTLCTIVNTVFEMAMEHYPTMDAFD-----ANLQAGNIVFTVFTWE 886
QY 289 LVTEGVCLITGLGFNGPFEKAKWDA--CANMKWLFETNPFRTIASF----- 337
Db 887 MAFK--IIPDPYYPQK--KNINIDCVIIVLSLELSASKGSLSVLSRLRLRVF 939
QY 338 -----NINTNAMVARYFKRLKFLGNKLSGSLFLAL----- 372
Db 940 KLAQSMPLNT--LIKIIINSVGLGNLTFLIATIVFIFALVGLQSLSEDYCKRKGVS 996
QY 373 -WHGLSGYLV-CFQMEFLIYVEROARLLIOESPPLSKLAATVLOPFYLVQOTIH-- 428
Db 997 VVNGSKRLRWMCDFHSEFLV-----FRILGGEIENM 1029
QY 429 WLFMGYSMTAFCLFTWDMKMLKYKSIYFLGHIFPLSLFIL 469
Db 1030 WYCMVESQKSLITL-----LFLTWMLGNLVANLFLAL 1063

RESULT 3
US-09-134-001C-4885
; Sequence 4885, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ. ID NOS: 5674
; SEQ. ID NO. 4885
; TYPE: PR1
; LENGTH: 412
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4885

Query Match 4.0%; Score 102.5; DB 4; Length 412;
Best Local Similarity 16.0%; Pred. No. 0.02;
Matches 76; Conservative 74; Mismatches 165; Indels 159; Gaps 17;

Db 220 -FLYKIIIVLLQTYAVNPLQMLHGFTH-----MPLYMAVSYLFFDFAGYSLFAIVS 274
QY 325 -----ETNPRFTGSIASFINNT--NMAVARYIF-----KRLKFGNKEISQG----- 364
Db 275 YLYGKITPPNFKOPFAKAKIKDWNRMHMTLSFWPDCIYMRSLFYMSRKULKSQFAMS 334
QY 365 -----LSLFLALMHGHSGLVLCFQMEPLIVEROAARLIOESPTLSKLAITVLOPF 419
Db 335 NNAVFLNPFIMGVHGLE-----V 353
QY 420 YIVVQOTIH-WLPMGISMATFCLFTWDKMLKVYKSYIFLGHIFLSLFLPIY 472
Db 354 YIVVGLVLYHAALPIGYGY-----YERMRKQHPRMQNGFTVAISIIIFHFV 400

RESULT 4
US-09-328-352-7069
Sequence 7069, Application US/09328352
Patent No. 6562958

GENERAL INFORMATION:
APPLICANT: Gary L. Beaton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7069
LENGTH: 443
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7069

Query Match
Best Local Similarity 21.3%; Score 99.5; DB 4; Length 443;
Matches 61; Conservative 42; Mismatches 90; Indels 93; Gaps 11;

QY 11 TVVALA--GVLOSGF-----OELSLNKLTATSGASEQA--LRLLISIFAGYFAL 56
Db 185 TVVALACIGLFRFLIPKVAAPDDIDMNFIRTSQTKYKASLPVLTILLGSMW-- 242
QY 57 FYRHYLFYKETYL-----IHLEHTPTGLSIAYFNFGNQLYHSLCTIVLOFLRL 106
Db 243 ---YMFYGATVLTQIPQLTQONLHASENVVSLTLTFVSGIV--GSLCRKIGGSEINI 297
QY 107 MGRITIAV-LTTCFCQMAVLA-----GYTTATGNVDIKMTMPHCVLTKLIGLAVDY 159
Db 298 KWPVIGAVGLTVPALYLAAVLAVPERTGALLTKDVFQGWYVYHMAVLTLLGIS-- 354
QY 160 FDGKDQNSLSSEQKVAIRGVPSLLEVAGFSYFYGAFLVGPOFSM-----NHWKLV 212
Db 355 -----GGFYIVPLYAMQAVSPSHRAKV 379
QY 213 QGELIDIPKIPISIIIPALKRLSLGLFYLVGYTLSPHTEDEYLT 258
Db 380 AA-----NNILNAVFMVSSAIFSLILSLVKIDIKILFLIT 415

RESULT 5
US-09-352-991A-29870
Sequence 29870, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIORITY APPLICATION NUMBER: US 60/074,788
PRIORITY FILING DATE: 1998-02-18
PRIORITY APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 29870
LENGTH: 243
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29870

Query Match
Best Local Similarity 23.3%; Score 98.5; DB 4; Length 243;
Matches 65; Conservative 36; Mismatches 91; Indels 87; Gaps 13;

QY 225 NSIIPALKRLSLGLFYLVGYTLSP-----HITEDVLTEDYDNDHPFRCMYLWNGEV 280
Db 28 DGLLVTLKLISL--LIGLLAVPLALMRVSKQPLV-----NFPAM 67
QY 281 LYKYTCMLVTEGVCILTGLGNGPEKAKWDACANNKWLFEINPRFTGSIASFIN 340
Db 68 LVTY-----VIRGTPMLVQPLFIY---GLAOFDAVESALMPWLSNASFACLA-FAIN 118
QY 341 TNAMVARYIFKRLKFLGNKEL-----SQGLSLFLALMHGHSGLVLCFQMEPLIVERO 396
Db 119 TSYVTEILAGSLKATPHGEIEAKAMGMSRL-----KMYRIILPSAL 162
QY 397 AARLIOESPTLSKLAITVLOPEYVVOQTIHMLFMYSMTAFCLFTWDKMLKVYSIYF 456
Db 163 RRLPQYSNEVIMMQLDTSIASIVTLVDIT-----GAARTVY-----SOY 203
QY 457 L-----GHIFLSLFLPIYIHKR-----WYPRK 480
Db 204 LPPEAFITAGLFYLCFTLVRLFLKLAERRMLAYLAPRK 242

RESULT 6
US-09-107-532A-3878
Sequence 3878, Application US/09107532A
Patent No. 6581275

GENERAL INFORMATION:
APPLICANT: Lynn A. Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Daneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3878:
SEQUENCE CHARACTERISTICS:
LENGTH: 401 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...401
SEQUENCE DESCRIPTION: SEQ ID NO: 3878
US-09-107-532A-3878

Query Match 3.8%; Score 98.5; DB 4; Length 401;
Best Local Similarity 18.6%; Pred. No. 0.052;
Matches 75; Conservative 45; Mismatches 136; Indels 147; Gaps 17;

QY 60 HYLKYEKYLHLPHFTGSLIAYNFGNQLYHSLCTVLOPL-----LRLMGRIT--- 112
DB 5 HMPYAEFPYFLL-----LAIA-----LVPIVLSLIMEKRLPYQSAVTLTF 47
QY 113 -----AVLTFCFQMAVLAGYVYATGNVDIKWMPHCVLTKLIGL--AV 157
DB 48 LFVSGGNYWQAGALLIGYVLMQTLVWGYFYRQRKQGTGMPRAVILAIPLFTVY 107
QY 158 DYFDGKDQNSLSSEQKYAIRGVPSLIEVAGFSY----- 192
DB 108 PFPEHGK-----ASLIGFLGISYLTFRKSVQIMEIRDMIKTKYKPRY 150
QY 193 -----FYCAFVLP-----QFSNHHYMLVQGLIDIDPK--IPNSIILAKRLSLGLFYL 241
DB 151 IQFLFFPTISSGPDYRREREKQDHPVDRKVDLGGKVQINIFVGLFKFTIG--YY 208
QY 242 VGYTLSPHITEDYLLTEDYDNHP--FWFRGMYMLMGKPYLYKYVTCMLVTEGVCLITG 299
DB 209 FGVYLL--PYVERIAL-----QHGISALVGYMYVISMILFPDPAGSLAVAGTSYMG 261
QY 300 LGFNGFEKGAKKADCANMKVWLFEINPRTGTIASFNI-----NTNAVARYIF 350
DB 262 -----YQTPINFNKPELISNIKEFNMNMHTLSFWRDYYV 297
QY 351 KRL-----KELNGKELSGLS-----LLFLALMHGHSGLYV 382
DB 298 MRLMFPMKKKPFKSKIVMSNIGYFALFLMGIMHGLTWYIYV 340

RESULT 7

US-09-328-352-8050
Sequence 8050, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328.352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 8050
LENGTH: 488
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-8050

Query Match 3.8%; Score 97.5; DB 4; Length 488;
Best Local Similarity 23.4%; Pred. No. 0.089;
Matches 55; Conservative 32; Mismatches 101; Indels 47; Gaps 9;

QY 11 TVALAGYLOSGFELSLINKLATSIGASEQALRLIISIFLGYPPALFYRHLYFYKETYLI 70
DB 98 SLVGAALVLAIVAAIHNNOJLALFAQNEALTL-----AAGVGHAMSK-----V 142
QY 71 HLPHFTGSLIAYNFGNQLYHSLCTVLOPLILRLMGRITVAV---LTFPCQMAVYL 126
DB 143 HLFELFVCCFVGAITFTASV-----AYGKLAAKMAKKTISGAMVKEVQALIF-VAMLA 195

QY 127 AGYYTATGNVDIKWMPHCVLTKLIGLAVDPDGKDQNSLSSEQKYAIRGVPSLIE 186
DB 196 CGFFFTTGNWTAFWANTALALAFGWVWIAF---VGGD-----MPVVSLL- 239
QY 187 VAGFSYFYGAFVLPQFSNMHYMLVQGLIDIDPKIPNSII--PALKRLSLGLFV 240
DB 240 -----NSISGMAAGIGFTLENMMLIVAGSLVGSGALLSYIMCAMNRSIINVLV 290

RESULT 8

US-09-252-991A-32219
Sequence 32219, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32219
LENGTH: 545
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32219

Query Match 3.7%; Score 96.5; DB 4; Length 545;
Best Local Similarity 21.5%; Pred. No. 0.14;
Matches 62; Conservative 43; Mismatches 105; Indels 79; Gaps 18;

QY 221 GKIPNSIIPALKRLSLGLFYLVGYT-----LSPHITEDYLLTE--DYDNHPFW 267
DB 125 GRINNAF--AVGRRRWGMALVFPATTIANYIDRAALGVMPILAEKSWTAMDYANINFW 182
QY 268 FRCMY--MLIWGEV-----LYKYVTCMLVTEGV--CILTGLGFN-----GFEK 308
DB 183 FQVGYAIGFILOGFRIDKGVKRAFFLAVILMSIATGAHGLANTSAGFMVCRFILGLTE- 241
QY 309 GKAKMDACA--NMKYWLEINPRTGTIASFNINMNAVARYIFRLKELNGKELSGLSL 367
DB 242 -AANYPACVKTTRVW--PPAGERAVAT--GIFNAGTNV-----GAVTPALLP 284
QY 368 LFLALMHGHSGLYVCFQMEFLIYI-----VERQARLI--QESPTLSKLA 411
DB 285 LILSW--GMQAAFLCNSALGLVWLVPFWRMYVNPDEHPRVKQSELEYIQQDEPPATRV- 342
QY 412 AITVLOPFYLVQOQTIHMLF--MGYSMTAFCLFTWDMKLVY--KSIYFLG 458
DB 343 -----PFSQILRRRGTAFAFALAYSIAPVFWFYLWLPPFLNQOYGLG 385

RESULT 9

US-09-328-352-7176
Sequence 7176, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328.352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7176
LENGTH: 551
TYPE: PRT
ORGANISM: Acinetobacter baumannii

US-09-328-352-7176

Query Match 3.7%; Score 95; DB 4; Length 551;
 Best Local Similarity 22.3%; Pred. No. 0.2;
 Matches 55; Conservative 38; Mismatches 96; Indels 58; Gaps 11;

QY 257 LTEDYDNP-----WFCMYMLGWKFLYKYVTCL-----VTEGVCIITGLGPN 303
 DB 53 LTPRIDGPFVHKPPLHMLSMFELFG-----VHVWLRVLPVAGTLMVLGL--- 102
 QY 304 GFEEKGRKMDACANMKWLEETNPFVGTIASF--INTNAVARIYFKLKFGLKEL 361
 DB 103 -FLVKKHISNVAGLTIIATMLLFFGSGQYINHDLLTSWITISVLCFVDF---TI 157
 QY 362 SQGLSLFLALMHGLHS-----GYLVCFQWFLVIVEROARLIQESPTLSLAIT 414
 DB 158 SAKSIIFLGITAGAAFLSKGLIGILIPGMILPMLITTKWKKIPSLPLMLLPL 217
 QY 415 VLQPFYLVQ---QTIHMLFMGYSMTAFCLFTWDK--WLKVYSIYFLGHIFLSLIF 467
 DB 218 IVSPMLYVQSKVPQFLHYFFIDQGFNFSSKEFNKKQPM-----CFYLMILF 265
 QY 468 I--LPYI 472
 DB 266 VSFLPWL 272

RESULT 10

US-09-328-352-7092
 ; Sequence 7092, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328.352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 7092
 ; LENGTH: 308
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-7092

Query Match 3.6%; Score 93; DB 4; Length 308;
 Best Local Similarity 20.4%; Pred. No. 0.14;
 Matches 70; Conservative 49; Mismatches 114; Indels 110; Gaps 15;

QY 11 TVVALAGVLOSQFOLSLINKLATSLSGA-----SEQALRLIISIFIGYPALFYRYL 62
 DB 12 TLIGSALIMASWGVFKHITTAIGPDVGITLISLNLILI-IFRVNPKLISKYLL 70
 QY 63 F-----YKETYLIHLFHTFTGLSIAYFNGQLYHSLCTIVLOPLIRLMGRTITAV 114
 DB 71 ILGAILFLAYELCFPALAYSKTAQQAIEVSIYVNLMPSL-TVLAFLVIFRELKENVPII 128
 QY 115 LTPFCQAYLLAGYYTATGNYD-----IKTMTPCVUTL 150
 DB 129 L-----GLLISIGIIFQTGNGDFSIGRVVDNFHNSPLSYLAIGAIMAF-YCVILTK 182
 QY 151 KLIGLAVDYFDGQKNSLSSEQKYAIRGVPSLLEVAGSYFAGFLVQPGSNHMYK 210
 DB 183 KM-----GKQNPIS-----IFFLGV-----ALTMWK 205
 QY 211 LVQGLIDIPGKIPIISIPAL-KRLSLGLFY-----LVGYTLSPHITE-- 253
 DB 206 LIFSQVSLPSLDLNTIISLIAAIGLYANMIGIIGHNITMLVVASYFTPIISLL 265
 QY 254 DYLLTEDYDNPFRFCMMLIWGFVLAKKYVTCMLVLEGVCI 296
 DB 266 AMLVQLTELISLSEW-QGTAMVTAGSLI-----CWLSTNMAVI 301

RESULT 11

US-09-066-047-7
 ; Sequence 7, Application US/09066047A
 ; Patent No. 6306394
 ; GENERAL INFORMATION:
 ; APPLICANT: MURPHY, Cheryl
 ; STOREY, James
 ; BELTZ, Gerald A.
 ; COUGHLIN, Richard T.

TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND METHODS OF
 USE OF GRANDLOCYTIC ERLIICHA
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: HALE AND DORR LLP
 STREET: 60 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: United States
 ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/066,047A
 FILING DATE: 24-Apr-1998
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/044,869
 FILING DATE: 25-Apr-1997
 ATTORNEY/AGENT INFORMATION:

NAME: Superko, Colleen
 REGISTRATION NUMBER: 39,850
 REFERENCE/DOCKET NUMBER: 106,941.156
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 526-6000
 TELEFAX: (617) 526-5000
 INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 506 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-066-047-7

Query Match 3.6%; Score 93; DB 4; Length 506;
 Best Local Similarity 20.7%; Pred. No. 0.29;
 Matches 75; Conservative 45; Mismatches 133; Indels 110; Gaps 15;

QY 27 SLINKLATSLSGSEQALRLIISIFIGYPALFYRYLFLYKETYLIHLFHTFTGLSIAYPNF 86
 DB 164 SFGMLVSLPFTTKVALYVLKFPSEGSITLY-----FCIFTSIYAIRAF 208
 QY 87 GNQLYHSLCTIV-----LOFLIRLMGRTITAVLTPFCQMAI-LLAGYYTA---TG 135
 DB 209 LEQVVRRLMAYMFGQAGLMMAGCPISDILIVQLSFSVLVOLLGMFADSVVKRSK 268
 QY 136 NYDILKMTMPHVLVLTKLIGLAVDYFDGKQNSLSSEQKYAIRGVPSLLEVAGSYFYG 195
 DB 269 HVDIN-RMAGC--FKLASME-----AMGCIYALMLNGCFPTTAG 304
 QY 196 AFLVGPQPSNN-----HYMKLVQGLIDIPGKIPIISIPAL-KRLSLGLFYLVGYTL 246
 DB 305 FVTGMLMMLHMLQSFQVLYLKYQPMLGWL-----FASNGMKLFWL---AC 348
 QY 247 LSPHITEYLLTEDYDNPFRFCMMLIWGFVLAKKYVTCMLVTEGCIITGLGPNFE 306
 DB 349 LKPCST-----TPRYAPSPSSKLSIIML-----SLITVSGVLVYEGGLIFS--E 391

QY 307 EXGAKMDACANMKWLFETNPRFTGTIASFNINTNAVARYIPKRLKELGNKELSGIS 366
DB 392 HKFVTFGVATKTLWL-----CGVYLF-----FILPRRQPLGNYESAIGDS 433
QY 367 LIF 369
DB 434 WYV 436

RESULT 12

US-09-107-532A-5141
Sequence 5141, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: LYNN A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5141:

SEQUENCE CHARACTERISTICS:

LENGTH: 793 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...793
SEQUENCE DESCRIPTION: SEQ ID NO: 5141:
US-09-107-532A-5141

Query Match 3.6%; Score 93; DB 4; Length 793;
Best Local Similarity 20.8%; Pred. No. 0.57;
Matches 64; Conservative 47; Mismatches 115; Indels 82; Gaps 13;

QY 12 VVALAGVLOSROELSLKATSLGASDAQALIIISIFLGYFALFYRYLYFYKETYLIH 71
DB 335 IIVLSAVVYFAGIMVLLSTI-TNLSAISRLFOVLL-----PFS-----FNFLDQTVNL 382
QY 72 LFTFTGL-----SIAYFNG-----NOLYHSLGIVLQFLILRLMGRT- 110
DB 363 VGFLLGLARALSAVYKKAFLPTIALLSGILNNTTRILSWQLIIVYVILILAVWSRIE 442
QY 111 -----ITAVLTTFPCFQWAVYLLAGYY-----TATGNY-----DIKWTWP 144

DB 443 FYRTKFVYVSWGALIFDASLFGFLF-IYVIALGYHSGRFLVGRITGNHFLFPPSDVWPSG 501
QY 145 HCVLTLKILGLANDYFDGSKDNLSSEQOK-----YAIRVPSLLEVAGSYFY- 194
DB 502 LIGLISLIALVALYQYLAEDTTLKLGSEYKDRFDRLLLETGTATSHQLRLSGSYFY 561
QY 195 -----GAFVVGQPSFMNHYMKLVQVELIDIPKIPNSIIPALKRSLGLFVLYGTYLLSP 249
DB 562 QEDDDQVVFQYQIGNKC--FVLGD PVGNPQMMAATNQPFKQAD-----CLGQLAFY 614
QY 250 HITEDYLL 257
DB 615 KISESYV 622

RESULT 13

US-09-097-889-25
Sequence 25, Application US/09097889
Patent No. 6218117

GENERAL INFORMATION:

APPLICANT: Herinstdt, Corrina
APPLICANT: Ghosh, Soumitra S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/097,889
FILING DATE: 15-JUN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Roseman Ph.D., Stephen J.
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 660088.417
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-097-889-25

Query Match 3.6%; Score 92; DB 3; Length 380;
Best Local Similarity 20.1%; Pred. No. 0.24;
Matches 54; Conservative 32; Mismatches 75; Indels 108; Gaps 16;

QY 201 PQFSNMHYMKLVQVELIDIPKIPNSIIPALKRSLGLFVLYGTYLLSP 249
DB 3 FWRKINPLMKLINSFIDLPTPSNISAMWNFSLGACILQITTLFLAMHY---SP 57
QY 250 -----HITED-----YLLTEDYDNHPWPCWYM-----LIWGFVLYKYVT 286
DB 58 DASTAFSSIAHITRDVNYGAIIRYL--HANGASMFICFLHIGRGLYYSFL--YSE 111
QY 287 CMLVTEGVCIL---TGLQFNGFEEXGAKMDACANMKWLFETNPRFTGTIASFNINTNA 343

Db 112 TWNI--GIILLATWATRMFGY-----VLPW---GOMSEW-----GA 143
Qy 344 WVAIYFKRLKFNKLSQGLSLFLALMGLHSGYLVCFQMEFLIYVERQARLIQE 403
Db 144 TVTNLSAIPYIGTD-----LVQW---IMGYSV-----D 171
Qy 404 SPTLSKLAATVLOPFYLVQOTIHWLFM 432
Db 172 SPTLRPFTHFLIPFIATLHLFL 200

RESULT 14
US-09-098-079-25
Sequence 25, Application US/09098079
Patent No. 6489095
GENERAL INFORMATION:
APPLICANT: Herrnstadt, Corrina
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Cleveland, William
APPLICANT: Fahy, Boia F.
APPLICANT: Davis, Robert E.
TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF
TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098, 079
FILING DATE: 15-JUN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Roseman Ph.D., Stephen J.
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 660088, 416
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
TELEFAX: (206) 682-4900
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-098-079-25

Query Match 3.6%; Score 92; DB 4; Length 380;
Best Local Similarity 20.1%; Pred. No. 0.24;
Matches 54; Conservative 32; Mismatches 75; Indels 108; Gaps 16;

Qy 201 PQSMNHMKLVQGLIDIPKIPNSIIP-----AKRLSLGLFYLGVYTLSP 249
Db 3 PMRKINPLMKLINISFDLP--TPSNISAMNFGSLGACILITGLFLAMHY---SP 57
Qy 250 -----HITED-----YLTEDYDNHPFWRQMTM-----LIMGFVLKYKT 286
Db 58 DASTAFSSIAHTIDVNYGWIIRYL--HANGASMFYICLFLHIGRGLYGSFL--YSE 111
Qy 287 CMLTEGVCIU---TGLGFNGFEEKKAKMDACAMKVMLEFETNPRFGTIASFININTNA 343
Db 112 TWNI--GIILLATWATRMFGY-----VLPW---GOMSEW-----GA 143
Qy 344 WVAIYFKRLKFNKLSQGLSLFLALMGLHSGYLVCFQMEFLIYVERQARLIQE 403

Db 144 TVTNLSAIPYIGTD-----LVQW---IMGYSV-----D 171
Qy 404 SPTLSKLAATVLOPFYLVQOTIHWLFM 432
Db 172 SPTLRPFTHFLIPFIATLHLFL 200

RESULT 15
US-09-328-352-5397
Sequence 5397, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328, 352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 5397
LENGTH: 470
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5397

Query Match 3.6%; Score 92; DB 4; Length 470;
Best Local Similarity 21.2%; Pred. No. 0.33;
Matches 66; Conservative 46; Mismatches 124; Indels 76; Gaps 16;

Qy 138 DIKWTM--PHCVTLTKLIGLAVDFDGGKDQNSISEQKRYAIRGVPSLL-EVAGSFYFY 194
Db 73 DTSWILISTALVLMITIPGLALFYGGMVRKKNVLTWPFSLAAILVSLMLWVINGSIAP 132
Qy 195 GAVLVERQPMNHMKLVQGLID-IPKIPNSIIPALKRLSLGLFYLGVYTLSPHITE 253
Db 133 SG--TGAYFG-DISKAMLVGAPDALSCTIPESLFIQW---TFALITVALISGSA- 184
Qy 254 DYLTEDYDNHPFWRQMTMIMKGFVLKYKVCMLVTEGVCILTGLGFNGFEEKKAKW 313
Db 185 -----DMKKSAPFAFAIWM-VLVVAPITMV-----W 212
Qy 314 DACAMKVMLEFETNPR--FTGTIASFININTNAVARIYFKRLKFNKLSQ----- 363
Db 213 AADG---WLFKAGALDPAGGTVVHINSVAGLVAAYMLKRLIGLGRSMAPHLTLTVI 268
Qy 364 GLSLFLALMGLHSGYLV-----CFQMEFLIYVERQAA-----RLQESTL----- 407
Db 269 GASLLWVG-WFGFNGSALGAGARASMAILVTQVAAAAAASFMLVVERMIRGKASVIGA 327
Qy 408 -SKLAATVLOP 418
Db 328 SGAVAGLVITP 339

Search completed: November 21, 2003, 14:17:10
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:16:08 ; Search time 35 Seconds
(without alignments)
2540.185 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 2580
Sequence: 1 MASSAEGDEGTVALAGVLQ.....ILFYIKAMPKREKKE 487

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PC1_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PC10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2580	100.0	487	US-09-938-803-6	Sequence 6, Appl1
2	350	13.6	488	US-10-053-248-6	Sequence 6, Appl1
3	310	12.0	57	US-10-029-386-29742	Sequence 29742, A
4	231	9.0	425	US-09-821-812-9	Sequence 9, Appl1
5	215	8.3	58	US-09-864-761-33415	Sequence 33415, A
6	150	5.8	342	US-10-205-219-192	Sequence 192, App
7	123.5	4.8	141	US-09-821-812-7	Sequence 7, Appl1
8	111	4.3	392	US-10-287-274-375	Sequence 375, App
9	105	4.1	613	US-09-815-242-13958	Sequence 13958, A
10	101	3.9	263	US-09-769-787-119	Sequence 119, App
11	99.5	3.9	663	US-09-781-880-5	Sequence 5, Appl1
12	96.5	3.7	587	US-09-815-242-5694	Sequence 5694, App
13	96.5	3.7	603	US-09-815-242-12667	Sequence 12667, A
14	95.5	3.7	286	US-09-731-872-287	Sequence 287, App
15	95.5	3.7	286	US-09-731-872-287	Sequence 287, App

16	95.5	3.7	286	12	US-09-876-997-287	Sequence 287, App
17	93.5	3.6	613	9	US-09-815-242-10225	Sequence 10225, A
18	93.5	3.6	697	10	US-09-881-752A-76	Sequence 76, Appl1
19	93	3.6	819	10	US-09-796-720B-4	Sequence 4, Appl1
20	93	3.6	819	12	US-10-146-733-24	Sequence 24, Appl1
21	92.5	3.6	316	11	US-09-510-332-121	Sequence 121, App
22	92	3.6	380	9	US-09-098-079-25	Sequence 25, Appl1
23	91.5	3.5	473	15	US-10-000-151B-2	Sequence 2, Appl1
24	91	3.5	570	11	US-09-922-225A-12	Sequence 12, Appl1
25	91	3.5	627	15	US-10-090-455-8	Sequence 8, Appl1
26	91	3.5	646	12	US-10-079-087-2	Sequence 2, Appl1
27	91	3.5	646	14	US-10-072-621-9	Sequence 9, Appl1
28	91	3.5	646	14	US-10-154-452-4	Sequence 4, Appl1
29	91	3.5	646	14	US-10-154-452-8	Sequence 8, Appl1
30	91	3.5	646	15	US-10-090-455-2	Sequence 2, Appl1
31	90.5	3.5	403	10	US-09-738-626-3904	Sequence 3904, App
32	90.5	3.5	622	9	US-08-843-558-11	Sequence 11, Appl1
33	90.5	3.5	637	8	US-08-879-337-8	Sequence 8, Appl1
34	90.5	3.5	1564	10	US-09-801-368-244	Sequence 244, App
35	89.5	3.5	338	10	US-09-976-605-6	Sequence 6, Appl1
36	89	3.4	388	9	US-09-741-669-370	Sequence 370, App
37	89	3.4	388	9	US-09-912-020-378	Sequence 378, App
38	89	3.4	874	10	US-09-963-766-6	Sequence 6, Appl1
39	88.5	3.4	206	11	US-09-791-279-129	Sequence 129, App
40	88	3.4	305	11	US-09-510-332-135	Sequence 135, App
41	88	3.4	664	10	US-09-898-533-2	Sequence 2, Appl1
42	88	3.4	664	15	US-10-176-847-40	Sequence 40, Appl1
43	88	3.4	697	10	US-09-989-919-108	Sequence 108, App
44	88	3.4	697	11	US-09-919-039-200	Sequence 200, App
45	88	3.4	697	12	US-10-024-298A-77	Sequence 77, Appl1

ALIGNMENTS

RESULT 1
US-09-938-803-6
; Sequence 6, Application US/09938803
; Patent No. US20020076762A1
GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Reddy, Roopa
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Yang, Junning
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: FULL-LENGTH EXPRESSED GENETIC MARKERS
; FILE REFERENCE: PF-0695 US
; CURRENT APPLICATION NUMBER: US/09/938, 803
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/311, 894
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte Clone 1928920
US-09-938-803-6
Query Match 100.0%; Score 2580; DB 9; Length 487;
Best Local Similarity 100.0%; Pred. No. 8.9e-239;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASSAEGDEGTVALAGVLSGFOELSLNKATSLGSEQLRLITISFLCYPALFVRH 60
DB 1 MASSAEGDEGTVALAGVLSGFOELSLNKATSLGSEQLRLITISFLCYPALFVRH 60
QY 61 YLFYKETYLIHLFTFTGSLIAYFNFGNLVHSLCLIVLQFLILRLMGRTITAVLTTPCF 120

```

Db      61 YLFYKETYLIHLFHTFTGLSIAVFNFQNLHSLCLCTIVQFLRLMGRITAVLTTCF 120
      121 QMAVLLAGYYTATGYDICKMTMPHCVLTKLIGLVDPDGDKNSSSSROQXAIRG 180
      121 QMAVLLAGYYTATGYDICKMTMPHCVLTKLIGLVDPDGDKNSSSSROQXAIRG 180
      181 VPSLLEVAGSFYFGAFLVGPQPSNMNMYKLVOGELIDIPGKIPNSIIPALKRSLGLFY 240
      181 VPSLLEVAGSFYFGAFLVGPQPSNMNMYKLVOGELIDIPGKIPNSIIPALKRSLGLFY 240
      241 LVGYTLISPHITEDYLLTDDYDNNHPFPCMTMLNGKVLVLYKVTCTVTEGVCILTGL 300
      241 LVGYTLISPHITEDYLLTDDYDNNHPFPCMTMLNGKVLVLYKVTCTVTEGVCILTGL 300
      301 GENGFEKAKAMDACANNKMWLFETNPRFTGTIASFNINTAMVARYFKELKPLGNKE 360
      301 GENGFEKAKAMDACANNKMWLFETNPRFTGTIASFNINTAMVARYFKELKPLGNKE 360
      361 LSGSLSLFLALMHGHSGLVCFQMEFLIVIVERQARLIDESPPLSKLAITVLPQPY 420
      361 LSGSLSLFLALMHGHSGLVCFQMEFLIVIVERQARLIDESPPLSKLAITVLPQPY 420
      421 YLVQOTIHWLFNGYSMTAFCLFTWDMKLYKYSIYFLGHIPLSLFLIPYIHKANVPRK 480
      421 YLVQOTIHWLFNGYSMTAFCLFTWDMKLYKYSIYFLGHIPLSLFLIPYIHKANVPRK 480
      481 EKLKME 487
      481 EKLKME 487

```

```

RESULT 2
US-10-053-248-6
; Sequence 6, Application US/10053248
; Publication No. US20030144188A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Nucleic Acid
; FILE REFERENCE: P-IS 4814
; CURRENT APPLICATION NUMBER: US/10/053,248
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-248-6

```

```

Query Match      13.6%; Score 350; DB 12; Length 488;
Best Local Similarity 13.6%; Pred. No. 6.3e-25;
Matches 108; Conservative 82; Mismatches 177; Indels 76; Gaps 13;

      55 ALFYRHLYFYKST--YLHLFHTFTGLSIAVNFQNLHSLCLCTIVQFLRLMGRIT 112
      39 ALMPFRYLSSTKTSFIRHVAVTLGLVLAFCFGYALHPLVQSGISYICMIIGVE-- 96
      113 AVLTPFCPMAYLLAGYYTATGYDICKMTMPHCVLTKLIGLVDPDGDKNSSS 170
      97 -----NMNPMIITOKITSLACIHDGMRKDE-ELT 128
      171 SROOKAIRGVPSLLEVAGSFYFGAFLVGPQPSNMNMYKLVOGELIDIP----- 222
      129 SSORLAVRMPSLLEYLSYNCGMILAGPLCSYVDYTFIIGRSYHITOGGNGKEPT 188
      223 -----IPNSIIPALKRSLGLFLVGYTLSPHITEDYLLTDDYDNNHPF--FCQWYL 274
      189 QYTERPSPNTAV-VQKLVQGLSLFLHITICT-LPVSYNIDENHQAASPTKIIVY 246
      275 INGKFLVLYYVTCWLVTEGVCILTGLGNGFEKAKAMDACANNKMWLFETNPRFTGT 334

```

```

Db      247 ISLARPXYFAWTIADAINNAGGFRGYDENGARWDLISNLRIQIEMSTSEKML 306
      335 ASFNINTAMVARYFKELKPLGNKELSQSLFLALMHGHSGLVCFQMEFLIVIVE 394
      307 DWNMIQTALMKRVCEYERTSPPTQ-----TFILSALMHGVPYIYTLTGLVMTL-- 359
      395 ROARLIGES-----PLSKLAITVLPQPYLVQOTIHWLFNGYSMTAFCLFTWDMK 447
      360 --AARAMNNRPHRYTEPSQKL-----FYDVTIWTIVQVASTVTPFVLSIKRS 409
      448 LKVYSIYFLGHIPLSLFLIP 470
      410 LTFYSSWYCCILIGILVLLLP 432

```

```

RESULT 3
US-10-029-386-29742
; Sequence 29742, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29742
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

; OTHER INFORMATION: MAP TO CHR12.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.5
; OTHER INFORMATION: SWISSPROT HIT: P54336, EVALUATE 8.00e+00
US-10-029-386-29742

```

```

Query Match      12.0%; Score 310; DB 12; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.5e-22;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

      292 EGVCLTGLGNGFEKAKAMDACANNKMWLFETNPRFTGTIASFNINTAMVARY 347
      2 EGVCLTGLGNGFEKAKAMDACANNKMWLFETNPRFTGTIASFNINTAMVARY 57

RESULT 4
US-09-821-812-9
; Sequence 9, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; FILE REFERENCE: P-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens

```

FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(425)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-821-812-9

Query Match 9.0%; Score 231; DB 12; Length 425;
Best Local Similarity 25.2%; Pred. No. 1.3e-13;
Matches 69; Conservative 45; Mismatches 122; Indels 38; Gaps 7;

118 FCFQWAVL-----LAGYYTANGNYDKMTMPHCUTLTLGLAIVDFG--GKQNSLSS 171
11 FVFALGYTLVCQVRVYLFYDGYSDPSGPMIITOKITSLACIHGMFRKE-ELTS 69
172 EQQKAIAGVPSLELVAGFSYFYGAFLVGPQFSNMHYMKLVQGLIDIPGKIPTS----- 226
70 SQRLAVARMPSLLELYLSYNCNFMGLIAGPLCSYDYITFIEGRSYHTLQSGENKERTQ 129
227 -----IIPAL-----KRSLGLFVLGYTLLSPHITEDYLLTDEDYD 262
130 YERTVAISKCKVMRFIMSLYSWYXAXYQKLLVGLSLFHLITCT-LPVEYNIDHRQ 188
263 NHPRF-FCMTMLMGKVLVYVCULVTEGVCILTLGFGNGFEKAKADACANNKV 321
189 ATASPTKIIYLYLSLAAAPRYPAWTLADAINNAAGFRGYDENGAARDLISLRI 248
322 WLFETNPRFTGSIASFNTINAMVARYIFKRLKF 355
249 QQIENSTSFKMFLENDNMNIQTALMPKRCYERISF 282

RESULT 5
US-09-864-761-33415
Sequence 33415, Application US/09864761
Patent No. US20020048763A1

GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT FILING DATE: US/09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anncmax Sequence Listing Engine vers. 1.1
SEQ ID NO 33415
LENGTH: 58
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006512.12
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.6
OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.9
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
OTHER INFORMATION: EST HUMAN HIT: W26367.1, EVALUATE 3.00e-26
OTHER INFORMATION: SWISSPROT HIT: P43288, EVALUATE 4.50e+00
US-09-864-761-33415

Query Match 8.3%; Score 215; DB 9; Length 58;
Best Local Similarity 89.4%; Pred. No. 3.3e-13;
Matches 42; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

391 VIVERQARLIQESPTLSKAATVLOPFFYLVQOTIHWLMFGYSMT 437
12 LILPQARLIQESPTLSKLAITVLOPFFYLVQOTIHWLMFGYSMT 58

RESULT 6
US-10-205-219-192
Sequence 192, Application US/10205219
Publication No. US20030138803A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alistair
APPLICANT: Brooksbank, Robert
APPLICANT: Pinnock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: W-1-A-018200
CURRENT FILING DATE: US/10/205,219
PRIOR APPLICATION NUMBER: 2002-07-24
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 192
LENGTH: 342
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: BBI
US-10-205-219-192

Query Match 5.8%; Score 150; DB 12; Length 342;
Best Local Similarity 24.5%; Pred. No. 5.9e-06;
Matches 67; Conservative 35; Mismatches 121; Indels 50; Gaps 11;

226 SIIPALKRL-----SLGFLVGYTLLSPHITEDYLLTED-YDNHPRFRCMYML-INGKF 279
57 SLRPLLRAMPAPLPLGLLF-----LLSSHLPLEAVEDAFYARPLPARLFYMPVFFAF 111

Qy 280 VLKRYTCMLTEGVCILATGLGFMGEERKKA-----KWDACA 317
Db 112 RMRFFV-AMIAECGCIAGFGAYPAARAGAGGPTLCCPPSPSEKASLEVDYETIR 170
Qy 318 NKKWMLFFENRPTGTIASFNINTNMAVARYIFKRLKPKGNKLSGLSLPLALMNGH 377
Db 171 NIDCSTDCVVRDGMWMTVMQWMLAQYIKSNP-ARSYVLRATMTMLSAVWHGLH 229
Qy 378 SGVLVCFQMEFLIVIERQAAARLIQSPPLSKLAITVLOPFYVLVQTIHMLF-----MG 433
Db 230 PGYVLSFLIPLCLAAE---GRL--BSALGRHS-----PGQKAMDWVNHFLKRAVD 278
Qy 434 YSMTAFCLFTMDPKMLKVYSIYFLGHIFFLSLL 466
Db 279 YVCMGFEVLISLADTLRYMASIYFCIHFLALAL 311

RESULT 7
US-09-821-812-7
; Sequence 7, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Iln. BiaoYang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; FILE REFERENCE: P-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-812-7

Query Match 4.8%; Score 123.5; DB 12; Length 141;
Best Local Similarity 31.1%; Pred. No. 0.0063;
Matches 32; Conservative 19; Mismatches 45; Indels 7; Gaps 3;
Qy 118 FCFQMAVL---LAGYYTATGNVYDIKMTMPHCVLTKLIGLAVYFDG--GKQNSLSS 171
Db 11 FVFAIGYLVCCVTRVYFDYQYSADFSGPMWITTKITSLACEIHDMPKRD-ELTS 69
Qy 172 EQQXYAIRGVPSLLEVAGSFYGAFLVGPFSMNHMYMLVOG 214
Db 70 SGRDLAVRMPSSLLEYLSYNCNFMGLAGPLCSYNDYITFIISG 112

RESULT 8
US-10-287-274-375
; Sequence 375, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Foreysch, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERETO
; FILE REFERENCE: ELITRA 008DVI
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-375

Query Match 4.3%; Score 111; DB 12; Length 392;
Best Local Similarity 21.9%; Pred. No. 0.039;
Matches 93; Conservative 49; Mismatches 131; Indels 152; Gaps 22;
Qy 79 LSIATFNGNQLYHSLLCIVLOFLRLMGRTTAVLTFCQMAVLAGYYTA-TGNY 137
Db 72 LAMAVYVNRGOYHPL---VRPALIASLFG-----YSLGGSITIDVORY 113
Qy 138 DIKMTMPHCVLTKLIGLAVYFDGKQNSLSSEQ-----KVAIRGVPSLLEVAG 189
Db 114 ---WMLPYF-----YIRGHFNVSUVEFTAVCMITTYIGVAMLEFAPALFERLG 158
Qy 190 FSYFGAFLVGPQFSMNHMYMLVOGELIDIPKIPNSIIPALKRLSLGFLVAVGYTLSP 249
Db 159 WKV-----SLQRNKVMFFIILAGALLP 181
Qy 250 HITE---DYLLTEDYDNHPFW-----FRCMYMLWGKFLVLYKYTCMLVTEGVCIL 297
Db 182 TWHQSMGSLMISAGYKHPMQSYEMLPFLSLTAFTMGFSI-----VIFGSLVQ 233
Qy 298 TGLGFNGFEERKQAKWACAMKMWLFPETNPRFTGTIASFNINTNMAVARYIFKRLKFLG 357
Db 234 AGLRNGFDEKS-----LF---VKLTITIS-----VLALFIVLR-G 267
Qy 358 NKELSQGLSLFLALMNGHSGYLVCFQMEFLIV---IVERQAAARLIQSPPLSKLAIT 414
Db 268 ELIYRDKLSLAFAG-----DFYSVMFIEVLMLFPLVLRVAKLRNDRMLF-LSALS 320
Qy 415 VLQPFYVLVQOTIHLFMNGYSMTAF-----CLF-TWPKMLKV-YKSIYFLGHIFFLSL 465
Db 321 AL-----LGCATWR-LTVSLVAFNPGGYAYFPTEWELLISIGFVAIEICAYIVLRL 372
Qy 466 LFIPL 470
Db 373 LPILP 377

RESULT 9
US-10-202-824-6
; Sequence 6, Application US/10202824
; Publication No. US20030176648A1
; GENERAL INFORMATION:
; APPLICANT: Wood, John N.
; APPLICANT: Akopian, Armen N.
; TITLE OF INVENTION: Ion Channel
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESS: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike, P.O. Box 15437
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/202,824
; FILING DATE: 26-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,656
; FILING DATE: 24-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:

LENGTH: 2132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-202-824-6

Query Match 4.1%; Score 105; DB 12; Length 2132;
Best Local Similarity 19.4%; Pred. No. 1.4;
Matches 113; Conservative 74; Mismatches 194; Indels 200; Gaps 28;

QY 7 GDEGVVALAGVLQSGFQF-----LSLNKATSLGASEQALRL-----TISIF-----49
DB 565 GEEGQLGVPTGELTGAPEGPALDTTGQKSFSLAGYLNEPFAQRAMSVSMTSVIEL 624
QY 50 ---LGYPFAL--FRHYLFYK-----ETYLHLFH-----PTGSIAYFNQGLYSHL 94
DB 625 EESKLCPCPCPLISFQKYLIMECCPKMKKFKMALFELVTDPAELTIT-----672
QY 95 LCIVLQPLILRL---MGRITVAULT-----FCFQMAVLLAG---YYTATGNYDI 139
DB 673 LCIVNTVFAMMEHYPMTDAPDAMLQAGNIVFTVFETMEMAFKIAFDPTY-----FQK 727
QY 140 KWTMHCVL-TIKLIGLAVDFDGGKQNSLSSEQKTAIRGVPSELLEVAGSYFYGAFL 198
DB 728 KWNIFDCVIVVLSLELS-----ASKGSLSS-----VLSRLALDPTTGQKSFSLAGY 775
QY 199 VGPOFSNMHYMKLVQ-----GELIDIPGKI PMSII-----PALKRLSLGFLY 240
DB 776 LNEPFRARAMSVSIMTSVIELESKLCPCPLISFQKYLIMECCPKMKKFKMALFE 835
QY 241 LVGYTLSPHIT-----EDYLLTEDYDNHPFRCMYMLIWGKEFLYKYVTCW 288
DB 836 LVTDPEAFELTITLCIVNTVFAMMEHYPMTDAPD-----ANLQAGNIVFTVFETME 886
QY 289 LVTEVCILTLGIFGFEKKAQDA--CANMKWLEFETNPRFGTITASF-----337
DB 887 MAFK---IIPDPYFQK---KWNIFDCVIVVLSLELSASKGSLSVLSRLRLRVF 939
QY 338 ---NINTNVAARYIFKRLKFLGNKELSGSLFLAL-----372
DB 940 KLASWPLINT---IKTIIGNSVGLGNLFTLAIIVTFPALVGKQLSSEDCRCKDVS 996
QY 373 -WHGLSGYLVG-FQMEFLIVVERQAARLLOESPFLSKLAITVLQFPYLVQOTIH--428
DB 997 VMNGEKLAMHMDFFHSLV-----FRILGEMINM 1029
QY 429 WLENGYSMTAFCLFTWDMKLVKYSIYPLGHIFFLSLFL 469
DB 1030 WVCMEVSOKSICLI-----LFLTVMLGNLVLNLFIAL 1063

RESULT 10
US-09-815-242-13958
Sequence 13958, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13958
LENGTH: 613
TYPE: PRT
ORGANISM: Salmonella typhi
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(613)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13958

Query Match 3.9%; Score 101; DB 9; Length 613;
Best Local Similarity 17.1%; Pred. No. 0.64;
Matches 110; Conservative 83; Mismatches 185; Indels 264; Gaps 29;

QY 22 GFQSLNKL-----LATSIGASEQALRLIISIFLGFPALFYHYLYKET-----67
DB 14 GFVILASRGMSNLATIGVSGVGLAALVAVGDF--FANGKQAFSPLTWMSVG 71
QY 68 -----YLHLPHFE-----TSLIAYFNQGLYSHLSCI 97
DB 72 NFNIGFNLVLDGLSTMLSVTVGFLIHMFASTYMGEBYS-RFAYITLFLASNVVL 130
QY 98 VL-QFLIRLMGRTITAVLTTCFQMAVLLAGYIT-----132
DB 131 VLSNLLMLYMGVGVG-----CSYLLIGFYSDPKNGAAMKAFVTVRVGDVFLAF 183
QY 133 -----ATGNDIKRTMHCVLTILGILAVDFPFGKQDN 167
DB 184 ALFLYNELGLTNREWELEAPAFADGNMLMA-----TLMLLGAIV-----GKSAQ 232
QY 168 SLSEQKQVAIRG---VPSILEVAGSYFYGAFLVGPOFSNMHYMKLVQSGELIDIPGKI 224
DB 233 LPLQTLADAMAGTPPSALIHATM-VTAGVYLI---ARTHGLFMTPEILHLVG--284
QY 225 NSIIPALKRLSLGFLYV-----GYTLSPHITEDYLLTEDVD-----262
DB 285 --IIGAITLVWAGPALVQDIDKRVLAYSTMSQIGWFLA-----LGQAMDAALFHL 335
QY 263 -NHPFRCMYMLIWGKEFLV-----YK-----YVTGWLVBGVCIITLGIFN 303
DB 336 MTHAF-FKALFLASGVIILACHNEONI PKWGRLKSIPLVYACFLVGAALSLPLVTA 394
QY 304 GFEEKGAKMDACANMKWLEFETNPRFGTITASFINNTNVAARYIFKRLKFLNKLSEQ 363
DB 395 GFESKDELLAGANANGHI-----NLMVAGLVGAR--MTSLYTFPMIF-----434
QY 364 GLSLFLPALMG-----LHSGYLVCFQMEFLIVVERQAARLLOESPFLSKLAITVLQ 418
DB 435 -----IVFHGXKHIAHAGKXITHHLPLVLMI-----LSFVGAALVVP 474
QY 419 FYVLVQOTIH-----WLFMG-----YSMTAFCLFTW 444
DB 475 LQGLVPLQTELAGRVLTLEITSGVAIAGIXIAMIWMLGRVLTVTSIAXSAPCHLGTW 534
QY 445 DKMLKVKYSIYPLGHIFFLSLFLPYIRKAMVPRKXKLKM 486
DB 535 --WYNAM-----GFDWLYDKVFPKPLGIAMLLKRPDLNAL 568

RESULT 11

```
US-09-769-787-119
; Sequence 119, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 119
; LENGTH: 263
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-119

Query Match      3.9%; Score 99.5; DB 11; Length 263;
Best Local Similarity 21.6%; Pred. No. 0.29;
Matches 59; Conservative 40; Mismatches 99; Indels 75; Gaps 11;

QY 53 PALAFRHHLYFKETYLHLHFTFGSLAIYFNF-----GNQLYHSLLCIVLOF 101
DB 4 PLITFRHRKRVVLLTLFLALVYRDVLTFFFFDIHADLAKFDQALKNDLKSALDF 63
QY 102 LILRLMGRITVAVLTTCFQMAVLLAGYVYATAGNDIKMTMPCV-----LTLC-- 151
DB 64 RLQGFN-----LGRYQSFIIPIIIVLGGFQYELKXNVLSLGRVSYGKLKLTLOVA 119
QY 152 -----LIGLAVDYFDGCKDQ--NSLSSEQOKYALNGVPSLL--EVAGSFYFG 195
DB 120 SIPCLLYLVTLIAITTYFPGFSPDGNSLFSDDS-----GLQRLDGEIKSYLFFTC 174
QY 196 AFLVGPQSFNMHYMKVQSELIDIPKINSIIPAKRSLGL-----FVLGYT-- 245
DB 175 VLLIGIFNAIYFL-----QIVDVYGNVTRSAITTYMFLMSLMLYALPYMVPMTSL 229
QY 246 -----LISPHI--TEDYLLTPEYDNH 264
DB 230 MQASVGDVSLMKLFTYILYIVYMWLEKYEEN 262

RESULT 12
US-09-781-880-5
; Sequence 5, Application US/09781880
; Publication No. US20030166042A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Siles-Santiago, Immaculada
; TITLE OF INVENTION: No. US20030166042A1e1 Seven-Transmembrane
; TITLE OF INVENTION: Proteins/G-Protein Coupled Receptors
; FILE REFERENCE: 35800/208932
; CURRENT APPLICATION NUMBER: US/09/781,880
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/182,061
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 663
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-781-880-5

Query Match      3.9%; Score 99.5; DB 12; Length 663;
Best Local Similarity 21.1%; Pred. No. 0.99;
Matches 108; Conservative 62; Mismatches 196; Indels 147; Gaps 23;
```

```
QY 24 OELSLNKLTATSGASEQALRLI-----ISIFGYPALFYRHYLFYKERTYL--HLF 73
DB 116 EPLSNRRFTTALIGLVVCTICSCVMKTKQIMLSAHMLPLIARCLVPLETIYINKFA 175
QY 74 HTFTGLSIAYF-----NFGNQLYHSLLCIV---LOFLILRLMGRITVAVLTTCF 120
DB 176 MIFTLGLEVLYLGSNLLVPVNLAKSAYRELQVVEVYGLALGSLMNQVLVPVLFVFW 235
QY 121 QMAYLLAGYTTATGN-----YDIKMTMPCVLT--LILGLAVDYFDGCKDQNSLS 170
DB 236 LVLFLALQIYSYFSTEDQASREBRLIFLFTSIAECCSTPYSLGLVFTV----- 284
QY 171 SEQQYAIKRVPSLLEVAGSFYFAGFVGPFSNMHYMK-----LVQGLIDIPGK 222
DB 285 -----SFVALGVUTLCKF--YIQGRAPFNDP--AMNRGTEGVTILLIIVOTGLIEL--- 333
QY 223 IPNSIIPALKRSLSLGLFYLGVYTLSP--HITEDVLLTDEDVD--NHPFW--FR---CMYML 274
DB 334 ---QVHRAFLSLITLFIIVASIIQSMLEIDPVLVLAGSRDXSLWGHFRAVSLCLPFL 390
QY 275 IMGKVVLYKYVTCMLVTGVCILITGLGPNGEFEKAKADACANMKVLFETNRFCTGI 334
DB 391 VFPAYMAYM-----ICOFFHMDP-----WLL-----II 413
QY 335 ASFNINTNAAVARYIFKRLKFLGNKELSGSLFLALMGLHSGYLVCQFQMEFLIIVE 394
DB 414 ISSSLITSL-----QVLGTFLI-----YVFWMEERFKKEVE 445
QY 395 ROAARLIGESPTLSLAAITVLOPFYLVQOTI--HMLFMGYSMTAFCLFTMDKMLKYK 452
DB 446 NMDDVIYVNGTYRLLEFLVALCVVAVGVSEFTIRQEWTVWG-SMIIFHSYVNWLBRAQ- 503
QY 453 STYFLGHIFPLSLFILPYIKHAWPRKEXLKK 485
DB 504 -----LGWKSFLLRDAVANKIKSLPATEQOLEK 532

RESULT 13
US-09-815-242-5694
; Sequence 5694, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5694
```

LENGTH: 587
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5694

Query Match 3.7%; Score 96.5; DB 9; Length 587;
Best Local Similarity 21.0%; Pred. No. 1.6;
Matches 78; Conservative 52; Mismatches 138; Indels 103; Gaps 17;

QY 66 EYLLHLEHTFTGSLAVNFGNQLYHSLCTVQLFLRLMGRITTAVALTFECQMAVL 125
DB 40 DTFVIGSLITSLI-----SEYRTQKIDLEFWRRRK-RLIPAVLPICVLMFT 92
QY 126 L-----AGYATATGNVDIKMTMPHCVTLLKLGAVDPFGQKQNSLS 170
DB 93 LIFKPELLIQMRDAIAAIFVSNMWY-----ISQNVYFN----- 128
QY 171 SEQOKYAIRGVPSLLEVAGFSYFGAFLVGPQFSNNHYMKLVQGLIDIPKIPNSIIPA 230
DB 129 ---QFALEPRLKHLMSLAIEEQFYLLFPLVITFLHFRK-----PRNIQT 170
QY 231 L---KRLSLGFVLYGVYLLSPHITD---YLTFEDYDNHPFWRCMYMLIMGFVLYX 283
DB 171 LFIYSLISGLMIVHF-----ITGDSRYVFGT-DTRLQTLGLGCIAPFIMPEFALK 223
QY 284 YVTCMLVTEGVCILTGLGFGNFEKQKAKMDACANMKVLPETNPRF-----TGTIASF 337
DB 224 DISKIVS---LDIIGSGFAVL-MTLFIVGDQDMIT--NGFYIISFATLFIAT 276
QY 338 NINTNAAVARYI-FKRLKFLGNKELSQGLSLLFLALMH---GLHSGYLVCFQMEFLIYI 392
DB 277 AVHPSLSLPAKFLSMKPLIIGRYS-----LYLMHYPIIVFNYSYVQGIIVVYI 329
QY 393 VERQAKRLIQE 403
DB 330 IEILLTALMAE 340

RESULT 14

US-09-815-242-12667
Sequence 12667, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Heselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zykend, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12667

LENGTH: 603
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12667

Query Match 3.7%; Score 96.5; DB 9; Length 603;
Best Local Similarity 21.0%; Pred. No. 1.7;
Matches 78; Conservative 52; Mismatches 138; Indels 103; Gaps 17;

QY 66 EYLLHLEHTFTGSLAVNFGNQLYHSLCTVQLFLRLMGRITTAVALTFECQMAVL 125
DB 50 DTFVIGSLITSLI-----SEYRTQKIDLEFWRRRK-RLIPAVLPICVLMFT 102
QY 126 L-----AGYATATGNVDIKMTMPHCVTLLKLGAVDPFGQKQNSLS 170
DB 103 LIFKPELLIQMRDAIAAIFVSNMWY-----ISQNVYFN----- 138
QY 171 SEQOKYAIRGVPSLLEVAGFSYFGAFLVGPQFSNNHYMKLVQGLIDIPKIPNSIIPA 230
DB 139 ---QFALEPRLKHLMSLAIEEQFYLLFPLVITFLHFRK-----PRNIQT 180
QY 231 L---KRLSLGFVLYGVYLLSPHITD---YLTFEDYDNHPFWRCMYMLIMGFVLYX 283
DB 181 LFIYSLISGLMIVHF-----ITGDSRYVFGT-DTRLQTLGLGCIAPFIMPEFALK 233
QY 284 YVTCMLVTEGVCILTGLGFGNFEKQKAKMDACANMKVLPETNPRF-----TGTIASF 337
DB 224 DISKIVS---LDIIGSGFAVL-MTLFIVGDQDMIT--NGFYIISFATLFIAT 286
QY 338 NINTNAAVARYI-FKRLKFLGNKELSQGLSLLFLALMH---GLHSGYLVCFQMEFLIYI 392
DB 287 AVHPSLSLPAKFLSMKPLIIGRYS-----LYLMHYPIIVFNYSYVQGIIVVYI 339
QY 393 VERQAKRLIQE 403
DB 340 IEILLTALMAE 350

RESULT 15

US-09-731-872-287
Sequence 287, Application US/09731872
Patent No. US20020102604A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Joberc, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78-US3-REG
CURRENT APPLICATION NUMBER: US/09/731,872
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 287
LENGTH: 286
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -48...-1
US-09-731-872-287

Query Match 3.7%; Score 95.5; DB 10; Length 286;
Best Local Similarity 22.1%; Pred. No. 0.78;
Matches 76; Conservative 39; Mismatches 102; Indels 127; Gaps 19;

QY 161 DGGKQNSLSEQOKYAIRGVPSLLEVAGFSYFGAFLVGPQFSNN---HYMKLVQGLI 217
DB 6 DGGTSESLFDD---YASWGIRSTLMVAGFVYGLVGVCHQLSSLNATYRSLVAREKV 62

```

Qy 218 --DIPKIPNSIIPALKRLSLGLFVLVGYTLSPHITEDYLLTEDYDNHPFWRFCMYMLI 275
Db 63 FMDLAA---TRAIFGVQSTAGLMALLG-----DVLHAD-----KARGQN 101
Qy 276 WGFVLYKYVTCWLTVEGVCILTGLGPNGEKKAQKADACAMKVMLEFETNPRFTGTA 335
Db 102 W-----CWF-----HITATGFFCFE-----NVAVALSNLIFRTEDLFU---- 135
Qy 336 SFNINTNMAVARYIFRKLKFLG--NKELSQGLSLFLAL-----MHGLHSGYL 382
Db 136 -----VIHHLFAFLGLGCLVNLQGHYLAHTTLLLEMSPTTCVSMMLKAGW-- 184
Qy 383 CFQMEFLIVIVERQARLIQESPTLSKLAITVLQPFYLVQOTIHMLF-----MGYSMT 437
Db 185 -----SESL-----FWKLNQWLMTHMFHCRMVLYTHMW 212
Qy 438 AFCLFTWDKMLKVKYSIYFLGH--IFPLSL---LFILPY-IHK 474
Db 213 WVCFWHMDGLV---SSLYLPHLTLFLVGLALTLTLINPYWTHK 252

```

Search completed: November 21, 2003, 14:21:38
 Job time : 37 secs

RN WO200008155-A2.
 XX 17-FEB-2000.
 XX
 PF 06-AUG-1999; 99WO-US17777.
 XX
 PR 07-AUG-1998; 98US-0160065.
 PR 01-SEP-1998; 98US-0098703.
 XX
 PA (INCY-) INCYTE PHARM INC.
 PI Hillman JL, Yue H, Lal P, Tang YT, Gorgone GA, Guegler KJ;
 PI Corley NC, Baughn MR;
 XX WPI; 2000-205710/18.
 DR N-PSDB; AAF50889.
 XX
 PT New human receptor-associated proteins (HRAP) useful for the diagnosis,
 PT treatment and prevention of cell proliferative, autoimmune,
 PT inflammatory, reproductive, cardiovascular, and gastrointestinal
 PT disorders.
 PS
 XX Claim 1; Pages 73-74; 99pp; English.
 CC The present sequence is a human receptor-associated protein
 CC (HRAP) from Incyte clone 2906971 obtained from THYMOTUS CDNA library.
 CC This sequence is expressed in nervous, gastrointestinal
 CC and reproductive tissues. HRAP has cytostatic, immunomodulatory,
 CC antiinflammatory, cardiac, antiarteriosclerotic, hepatotropic,
 CC antiallergic, antirheumatic, osteopathic, antiallergic, antianemic,
 CC antiasthmatic, antidiabetic, dermatological and neuroprotective
 CC activities. The present sequence is useful in the diagnosis, treatment
 CC and prevention of disorders associated with HRAP expression, especially
 CC cell proliferative, autoimmune/inflammatory, reproductive,
 CC cardiovascular and gastrointestinal disorders (e.g. atherosclerosis,
 CC cirrhosis, leukaemia, cancer, AIDS, arthritis, allergies, anaemia,
 CC asthma, dermatitis, diabetes, osteoporosis, multiple sclerosis and
 CC irritable bowel syndrome).
 XX
 XX Sequence 487 AA;
 SQ
 Query Match 100.0%; Score 487; DB 21; Length 487;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASSAGDEGTVALAGVLQSGFQELSLNKLATSLGASDAQRLIISIFLGYFALFYH 60
 DB 1 MASSAGDEGTVALAGVLQSGFQELSLNKLATSLGASDAQRLIISIFLGYFALFYH 60
 QY 61 YLFYKETYLIHLFHTTGTSLAFNFGNQLYHSLCTIVLOFLRLMGRITAVLTTCF 120
 DB 61 YLFYKETYLIHLFHTTGTSLAFNFGNQLYHSLCTIVLOFLRLMGRITAVLTTCF 120
 QY 121 QMAVYLAGYYTATGNYDIKMTWPHCVLTKLGLAVDPDGKDNLSSEQOKXAING 180
 DB 121 QMAVYLAGYYTATGNYDIKMTWPHCVLTKLGLAVDPDGKDNLSSEQOKXAING 180
 QY 121 QMAVYLAGYYTATGNYDIKMTWPHCVLTKLGLAVDPDGKDNLSSEQOKXAING 180
 DB 121 QMAVYLAGYYTATGNYDIKMTWPHCVLTKLGLAVDPDGKDNLSSEQOKXAING 180
 QY 181 VPSLLEVAGFSYFYGAFLVGPQPSNMHMYKLVGELIDIPGKIPNSIIPALKRLSLGLFY 240
 DB 181 VPSLLEVAGFSYFYGAFLVGPQPSNMHMYKLVGELIDIPGKIPNSIIPALKRLSLGLFY 240
 QY 241 LVGYTTLSPHITEDYLLTDEYDNHPRFRCMWMMLNGKFLYKYVTCMLTBEVCILTGL 300
 DB 241 LVGYTTLSPHITEDYLLTDEYDNHPRFRCMWMMLNGKFLYKYVTCMLTBEVCILTGL 300
 QY 301 GFNGFEKGAKKADACANMKVMFLFETNPRTGTIASFNITNMAVARYIFKRLKFLGNKE 360
 DB 301 GFNGFEKGAKKADACANMKVMFLFETNPRTGTIASFNITNMAVARYIFKRLKFLGNKE 360
 QY 361 LSGGLSLFLALWHGLHSGYLVCFQMEFLIVIVEROAARLIQSPSLSKLAITVLOPFY 420
 DB 361 LSGGLSLFLALWHGLHSGYLVCFQMEFLIVIVEROAARLIQSPSLSKLAITVLOPFY 420

QY 421 YLVQOTIHLFWGYSMTAFCLFTMDKMLKVYKSYIFLGHIFSLFLIPYTHKAMVPRK 480
 DB 421 YLVQOTIHLFWGYSMTAFCLFTMDKMLKVYKSYIFLGHIFSLFLIPYTHKAMVPRK 480
 QY 481 EKLKKME 487
 DB 481 EKLKKME 487
 RESULT 2
 AAB31669
 ID AAB31669 standard; Protein; 487 AA.
 AC AAB31669;
 DT 30-APR-2001 (first entry)
 DE Amino acid sequence of a human protein having a hydrophobic domain.
 XX Human; hydrophobic protein; secretory protein; membrane protein; sepsis;
 XX tumour inhibition; immune deficiency; autoimmune disorder; anaemia; burn;
 XX infectious disease; cancer; ulcer; periodontal disease; coagulation;
 XX Parkinson's disease; fertility; immune response; thrombosis.
 OS Homo sapiens.
 XX
 PN WO200104297-A2.
 PD 18-JAN-2001.
 PF 16-JUN-2000; 2000WO-UP03942.
 PR 08-JUL-1999; 99UP-0194359.
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 PI Kato S, Kimura T;
 DR WPI; 2001-103081/11.
 DR N-PSDB; AAF25159, AAF25169.
 PT Isolated human proteins and polynucleotides are used in research and
 PT have activities including cell proliferation/differentiation activity,
 PT immune stimulating activity and receptor/ligand activity.
 PS
 XX Claim 1; Page 90-93; 151pp; English.
 CC The present sequence represents a human protein with hydrophobic domains.
 CC The protein possesses a hydrophobic domain and so is a secretory protein
 CC or a membrane protein. The protein is used as an antigen to prepare
 CC antibodies. The polynucleotide sequence is useful as a source of probes
 CC for genetic diagnosis. It is also useful for producing the protein
 CC in large quantities and for gene therapy. The eukaryotic cells are used
 CC for detecting the receptors or ligands corresponding to the protein and
 CC for detecting small novel pharmaceuticals. The antibodies are also used
 CC for detection, quantification and purification of the proteins. Both the
 CC protein and polynucleotide may be used in research or as nutritional
 CC sources or supplements. The protein may have cytokine and cell
 CC proliferation/differentiation activity, immune stimulating and cell
 CC activity, hematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity and tumour inhibition activity. It may therefore be used to
 CC treat immune deficiencies resulting from autoimmune disorders or
 CC infectious diseases, cancer, sepsis, anaemias, burns and ulcers,
 CC periodontal disease, Parkinson's disease, induce fertility, improve
 CC immune response and enhance coagulation or inhibit thrombosis.
 XX
 SQ Sequence 487 AA;
 Query Match 100.0%; Score 487; DB 22; Length 487;
 Best Local Similarity 100.0%; Pred. No. 0;

QY 481 EKLKME 487
 DB 481 EKLKME 487
 RESULT 4
 ABG06198
 ID ABG06198 standard; Protein; 424 AA.
 AC ABG06198;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #6189.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO2001/5067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS70385.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 36557; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II), or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pot_sequences.
 CC
 XX
 SQ Sequence 424 AA;
 Query Match 62.2%; Score 303; DB 22; Length 424;
 Best Local Similarity 100.0%; Pred. No. 5.1e-282;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 64 YKETYLILH.FHFTGTSLAIFNFGNQLYHSLLCTVLOGLILRLMGRITTAULTFPCFOMA 123
 |||||
 |||||

DB 1 YKETYLILH.FHFTGTSLAIFNFGNQLYHSLLCTVLOGLILRLMGRITTAULTFPCFOMA 60
 QY 124 YLLAGYTYTATGNVDIKWTMPHCVLTLLKILGLAVYFDGKDQNSLSSEQKXAIRGVPS 183
 |||||
 DB 61 YLLAGYTYTATGNVDIKWTMPHCVLTLLKILGLAVYFDGKDQNSLSSEQKXAIRGVPS 120
 QY 184 LLEVAGFSYFVGAELVGFQFSNMNHYMKLVQGLIDIPKIPISIIIPALKRISLGFYLVG 243
 |||||
 DB 121 LLEVAGFSYFVGAELVGFQFSNMNHYMKLVQGLIDIPKIPISIIIPALKRISLGFYLVG 180
 QY 244 YTLSPHITEDYLTLEDYDNHFPFRCMMLMGKFLVLYKYTCMLVTEGVCILTGLGFN 303
 |||||
 DB 181 YTLSPHITEDYLTLEDYDNHFPFRCMMLMGKFLVLYKYTCMLVTEGVCILTGLGFN 240
 QY 304 GFEKQKAKMDACANMKYWLFTNRPFTGTASFNINTNNAVARYIPKRLKLNKEISQ 363
 |||||
 DB 241 GFEKQKAKMDACANMKYWLFTNRPFTGTASFNINTNNAVARYIPKRLKLNKEISQ 300
 QY 364 GLS 366
 |||||
 DB 301 GLS 303
 RESULT 5
 AAG00235
 ID AAG00235 standard; Protein; 128 AA.
 AC AAG00235;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 4316.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 PI Dumas Malne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC00241.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT
 XX
 PS Claim 13; SEQ ID 4316; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 CC
 XX

SQ Sequence 128 AA;
 Query Match 25.3%; Score 123; DB 21; Length 128;
 Best Local Similarity 100.0%; Pred. No. 9.8e-110;
 Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MASSAEGDEGTVVALAGVLSQSGFQELSLNKLTATSLGASEQALRLIISIFLGPALFPRH 60
 DB 1 MASSAEGDEGTVVALAGVLSQSGFQELSLNKLTATSLGASEQALRLIISIFLGPALFPRH 60
 OY 61 YLFYKETLILHFFHTGSLAYENFGNQLYHSLLCIVLQFLILMGRITTAULTTCF 120
 DB 61 YLFYKETLILHFFHTGSLAYENFGNQLYHSLLCIVLQFLILMGRITTAULTTCF 120
 OY 121 QMA 123
 DB 121 QMA 123

RESULT 6
 ABG47472
 ID ABG47472 standard; Peptide; 58 AA.
 XX ABG47472;
 AC
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID NO 26120.
 XX
 KM Human, liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KM hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 XX
 PS Claim 27; SEQ ID NO 26120; 658bp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 1109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (1) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABG47348-ABG5930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 58 AA;
 Query Match 8.6%; Score 42; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.7e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 396 QARLIQSPFLSKAAITVLAQPPYLVQQTTHMLFMGYSMT 437
 DB 17 QARLIQSPFLSKAAITVLAQPPYLVQQTTHMLFMGYSMT 58

RESULT 7
 ABB27468
 ID ABB27468 standard; Peptide; 58 AA.
 XX
 AC ABB27468;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human peptide #119 encoded by breast cell single exon nucleic acid probe.
 XX
 KM Human, microarray; single exon probe; gene expression; breast;
 KM disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157271-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00662.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-496933/54.
 XX
 PT New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 XX
 PS Claim 27; SEQ ID NO 10436; 327bp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 58 AA;
 Query Match 8.6%; Score 42; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.7e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 396 QAARLIOESPTLSKLAITVLOPFYLVQQTTHLMFMGYSMT 437
 DB 17 QAARLIOESPTLSKLAITVLOPFYLVQQTTHLMFMGYSMT 58
 RESULT 8
 ABB32619
 ID ABB32619 standard; Peptide; 58 AA.
 AC ABB32619;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #125 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PS Claim 27; SEQ ID NO 25254; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 58 AA;
 Query Match 8.6%; Score 42; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.7e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 396 QAARLIOESPTLSKLAITVLOPFYLVQQTTHLMFMGYSMT 437
 DB 17 QAARLIOESPTLSKLAITVLOPFYLVQQTTHLMFMGYSMT 58
 RESULT 9
 ABB18117

ID ABB18117 standard; Protein; 58 AA.
 XX
 AC ABB18117;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Protein #116 encoded by probe for measuring heart cell gene expression.
 XX
 KW Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 PS Claim 15; SEQ ID NO 19887; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 58 AA;
 Query Match 8.6%; Score 42; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.7e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 396 QAARLIOESPTLSKLAITVLOPFYLVQQTTHLMFMGYSMT 437
 DB 17 QAARLIOESPTLSKLAITVLOPFYLVQQTTHLMFMGYSMT 58
 RESULT 10
 AAM53449
 ID AAM53449 standard; Protein; 58 AA.
 XX
 AC AAM53449;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 25554.
 XX

KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 OS Homo sapiens.
 FN WO200157275-A2.
 XX
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 25554; 650pp + Sequence listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 58 AA;
 Query Match 8.6%; Score 42; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.7e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 396 QAAALIOESPTSLKLAITVLOPFYVVOQTTHLWFMGYSMT 437
 DB 17 QAAALIOESPTSLKLAITVLOPFYVVOQTTHLWFMGYSMT 58
 RESULT 11
 ID AAM65826 standard; Protein; 58 AA.
 XX
 AC AAM65826;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 26132.
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 26132; 658bp + Sequence listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 SQ Sequence 58 AA;
 Query Match 8.6%; Score 42; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.7e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 396 QAAALIOESPTSLKLAITVLOPFYVVOQTTHLWFMGYSMT 437
 DB 17 QAAALIOESPTSLKLAITVLOPFYVVOQTTHLWFMGYSMT 58
 RESULT 12
 ID AAM13688 standard; Protein; 56 AA.
 XX
 AC AAM13688;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #122 encoded by probe for measuring cervical gene expression.
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -

PS Claim 27; SEQ ID No 18514; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes
CC (SENP: see A1110068-A1128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 58 AA;

Query Match 8.6%; Score 42; DB 22; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 396 QAARLIQESPTLSKLAITVLOPFYVVOQTTHLMFMGYSMT 437
Db 17 QAARLIQESPTLSKLAITVLOPFYVVOQTTHLMFMGYSMT 58

RESULT 13

AA026088
ID AAM26088 standard; Protein; 58 AA.

AC AAM26088;
XX

DT 17-OCT-2001 (first entry)
XX

DE Peptide #125 encoded by probe for measuring placental gene expression.
XX

KW Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder.
XX

OS Homo sapiens.
OS

PN WO200157272-A2.
XX

PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00663.
XX

PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX

PT WPI; 2001-488897/53.
XX

DR Human genome-derived single exon nucleic acid probes useful for
XX
XX analyzing gene expression in human placenta -
XX

PS Claim 27; SEQ ID No 26357; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP:
CC see A111315-A1157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX

SQ Sequence 58 AA;

Query Match 8.6%; Score 42; DB 22; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 396 QAARLIQESPTLSKLAITVLOPFYVVOQTTHLMFMGYSMT 437
Db 17 QAARLIQESPTLSKLAITVLOPFYVVOQTTHLMFMGYSMT 58

RESULT 14

AA001438
ID AAM01438 standard; Protein; 58 AA.

AC AAM01438;
XX

DT 09-OCT-2001 (first entry)
XX

DE Peptide #120 encoded by probe for measuring human breast gene expression.
XX

KW Probe; human; breast disease; breast cancer; development disorder;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX

OS Homo sapiens.
OS

PN WO200157270-A2.
XX

PD 09-AUG-2001.
XX

PF 29-JAN-2001; 2001WO-US00661.
XX

PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX

PT WPI; 2001-476286/51.
XX

DR Novel single exon nucleic acid probe used to measuring gene expression
XX
XX in a human breast -
XX

PS Claim 27; SEQ ID No 10178; 322bp; English.

XX The present invention relates to novel single exon nucleic acid probes
CC (see A1100010-A1110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridizes at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 58 AA;

Query Match 8.6%; Score 42; DB 22; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 396 QAARLIQESPTLSKLAITVLOPFYVVOQTTHLMFMGYSMT 437
Db 17 QAARLIQESPTLSKLAITVLOPFYVVOQTTHLMFMGYSMT 58

RESULT 15
 ABG35460
 ID ABG35460 standard; Peptide: 58 AA.
 AC ABG35460;
 XX
 DT 19-AUG-2002 (first entry)
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 25125.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 25125; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung; comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 58 AA:
 Query Match 8.6%; Score 42; DB 23; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.7e-32; Indels 0; Gaps 0;
 Matches 42; Conservative -0; Mismatches 0;
 Qy 396 QARLIQESPTLSKLAITVLPFYVVOQTTHWLFNGYSWT 437
 DB 17 QARLIQESPTLSKLAITVLPFYVVOQTTHWLFNGYSWT 58

Search completed: November 21, 2003, 14:22:36
 Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:20:54 ; Search time 20 Seconds

(without alignments)
2341.707 Million cell updates/sec

Title: US-09-938-803-6

Sequence: 1 MASSAEGDEGTVALAGVLQ.....ILPYHKAVPRKKLKME 487

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.8	556	2 A32466	numb protein - fru
2	9	1.8	1486	2 E96608	probable retroelasm
3	8	1.6	86	1 NZPG	endozeptine - pig
4	8	1.6	86	2 S63593	acyl-coenzyme A-bi
5	8	1.6	86	2 S63594	acyl-coenzyme A-bi
6	8	1.6	86	2 S63595	acyl-coenzyme A-bi
7	8	1.6	87	1 NZBO	endozeptine - bovin
8	8	1.6	88	2 A57711	diazepam-binding i
9	8	1.6	380	2 C84710	hypothetical prote
10	8	1.6	504	2 S57589	probable membrane
11	8	1.6	508	1 A43713	calcium-dependent
12	8	1.6	547	2 D95337	probable mangnese
13	7	1.4	54	2 H97854	hypothetical prote
14	7	1.4	55	2 D95140	hypothetical prote
15	7	1.4	86	2 S63592	hypothetical prote
16	7	1.4	87	1 NZRT	endozeptine [valida
17	7	1.4	87	1 NZRT	endozeptine - rat
18	7	1.4	87	2 A60059	endozeptine - mouse
19	7	1.4	100	2 B98008	hypothetical prote
20	7	1.4	101	4 S59321	hypothetical prote
21	7	1.4	107	2 S64345	hypothetical prote
22	7	1.4	126	2 S54159	probable sodium po
23	7	1.4	134	2 H84947	NADH2 dehydrogenas
24	7	1.4	150	2 G84173	archaeal histone A
25	7	1.4	152	2 E82875	hypothetical prote
26	7	1.4	162	2 AG0336	probable mannose-r
27	7	1.4	188	2 F81808	hypothetical prote
28	7	1.4	194	2 G83915	hypothetical prote
29	7	1.4	198	2 AF1569	hypothetical prote

30	7	1.4	228	2 F71831	hypothetical prote
31	7	1.4	242	1 S65576	phosphate uptake r
32	7	1.4	255	2 AE2740	outer membrane pro
33	7	1.4	256	2 T43869	phosphate transpor
34	7	1.4	259	2 C97521	27K outer membrane
35	7	1.4	260	2 T26496	coat protein - tom
36	7	1.4	268	2 T27223	hypothetical prote
37	7	1.4	272	2 T27223	hypothetical prote
38	7	1.4	272	2 D64933	hypothetical prote
39	7	1.4	275	2 D64933	NAD synthetase (EC 6
40	7	1.4	275	2 F90934	NAD synthetase (Im
41	7	1.4	275	2 B85783	NAD synthetase, pr
42	7	1.4	275	2 AC0709	NH3-dependent NAD
43	7	1.4	276	2 G72292	glycerol-3-phospha
44	7	1.4	276	2 AH0460	conserved hypot
45	7	1.4	292	2 T48327	hypothetical prote

ALIGNMENTS

RESULT 1

A32466 numb protein - fruit fly (*Drosophila melanogaster*)

C/Species: *Drosophila melanogaster*

C/Date: 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 24-Sep-1998

C/Accession: A32466

R/Jemura, T.; Shepherd, S.; Ackerman, L.; Jan, L.Y.; Jan, Y.N.

Cell 58, 349-360, 1989

A/Title: numb, a gene required in determination of cell fate during sensory organ forma

A/Reference number: A32466; MUID:89324081; PMID:2752427

A/Accession: A32466

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-556 <UEM>

A/Cross-references: GB:M27815; NID:G158000; PID:G158001

C/Genetics:

A/Genes: FlyBase:numb

A/Cross-references: FlyBase:FBgn002973

Query Match 1.8%; Score 9; DB 2; Length 556;

Best Local Similarity 100.0%; Pred. No. 1.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 360 ELISQSL 368

DB 369 ELISQSL 377

RESULT 2

E96608 Probable retroelement polypeptide P25P12.89 [Imported] - *Arabidopsis thaliana*

C/Species: *Arabidopsis thaliana* (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C/Accession: E96608

R/Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.R.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,

ansen, N.F.; Hughes, B.; Huizer, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

A./; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurov, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Xu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A/Reference number: AB6141; MUID:21016719; PMID:11130712

A/Accession: E96608

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1486 <STO>

A/Cross-references: GB:AE005173; NID:G9954746; PIDN:AA09097.1; GSPDB:GN00141

C/Genetics:

A/Genes: P25P12.89

A:Map position: 1
C:Superfamily: retrovirus-related polypeptide

Query Match 1.6%; Score 8; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SIFLYPFA 55
Db 752 SIFLYPFA 760

RESULT 3

NZPG
endoneptine - pig

N:Alternate names: diazepam-binding inhibitor
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 26-Feb-1999

C:Accession: S00805; S36839

R:Chen, Z.W.; Aggerberth, B.; Gell, K.; Andersson, M.; Mutt, V.; Oestenson, C.G.; Efendic
Burr, J. Biochem. 174, 229-245, 1998

A:Title: Isolation and characterization of porcine diazepam-binding inhibitor, a polypep-
tide release.

A:Reference number: S00805; MUID:88254787; PMID:3289918

A:Accession: S00805

A:Molecule type: protein

A:Residues: 1-86 <CHE>

R:Aggerberth, B.; Boman, A.; Andersson, M.; Joernvall, H.; Mutt, V.; Boman, H.G.
Eur. J. Biochem. 216, 623-629, 1993

A:Title: Isolation of three antibacterial peptides from pig intestine: gastric inhibitor
peptide.

A:Reference number: S36839; MUID:93387315; PMID:8375398

A:Accession: S36839

A:Molecule type: protein

A:Residues: 33-86 <ACE>

C:Comment: This protein may be the endogenous ligand of the benzodiazepine receptor.

C:Superfamily: endoneptine; endoneptine homology

C:Keywords: acyl-CoA binding; antibacterial; blocked amino end; intestine

F:1-86/Product: endoneptine #status experimental <MAT1>

F:33-86/Product: endoneptine (33-86) #status experimental <MAT2>

F:1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental

Query Match 1.6%; Score 8; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGRKMDA 315
Db 50 KGRKMDA 57

RESULT 4

S63593
acyl-coenzyme A-binding protein - turtle

C:Species: Testudines gen. sp. (turtle)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-1999

C:Accession: S63593

R:Kragelund, B.B.; Hojrup, P.; Skou Jensen, M.; Karlskov Schjerling, C.; Juul, E.; Knud
J. Mol. Biol. 256, 187-200, 1996

A:Title: Fast and one-step folding of closely and distantly related homologous proteins

A:Reference number: S63592; MUID:96190960; PMID:8609609

A:Accession: S63593

A:Molecule type: protein

A:Residues: 1-86 <KRA>

C:Superfamily: endoneptine; endoneptine homology

C:Keywords: acyl-CoA binding

F:18-74/Domain: endoneptine homology <NDZ>

Db 50 KGRKMDA 57

Query Match 1.6%; Score 8; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGRKMDA 315
Db 50 KGRKMDA 57

RESULT 5

S63594
acyl-coenzyme A-binding protein - mallard

C:Species: Anas platyrhynchos (mallard)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999

C:Accession: S63594

R:Kragelund, B.B.; Hojrup, P.; Skou Jensen, M.; Karlskov Schjerling, C.; Juul, E.; Knud
J. Mol. Biol. 256, 187-200, 1996

A:Title: Fast and one-step folding of closely and distantly related homologous proteins

A:Reference number: S63592; MUID:96190960; PMID:8609609

A:Accession: S63594

A:Molecule type: protein

A:Residues: 1-86 <KRA>

C:Superfamily: endoneptine; endoneptine homology

C:Keywords: acyl-CoA binding

F:18-74/Domain: endoneptine homology <NDZ>

Query Match 1.6%; Score 8; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGRKMDA 315
Db 50 KGRKMDA 57

RESULT 6

S63595
acyl-coenzyme A-binding protein - chicken

C:Species: Gallus gallus (chicken)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999

C:Accession: S63595

R:Kragelund, B.B.; Hojrup, P.; Skou Jensen, M.; Karlskov Schjerling, C.; Juul, E.; Knud
J. Mol. Biol. 256, 187-200, 1996

A:Title: Fast and one-step folding of closely and distantly related homologous proteins

A:Reference number: S63592; MUID:96190960; PMID:8609609

A:Accession: S63595

A:Molecule type: protein

A:Residues: 1-86 <KRA>

C:Superfamily: endoneptine; endoneptine homology

C:Keywords: acyl-CoA binding

F:18-74/Domain: endoneptine homology <NDZ>

Query Match 1.6%; Score 8; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGRKMDA 315
Db 50 KGRKMDA 57

RESULT 7

NZBO
endoneptine - bovine

N:Alternate names: acyl-CoA-binding protein; diazepam-binding inhibitor
N:Containing: endoneptine, short form
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1988 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999

C:Accession: A26448; S23127; A27886; A32944
R:Webb, N.R.; Rose, T.M.; Malik, N.; Marguardt, H.; Shoyab, M.; Todaro, G.J.; Lee, D.C.
DNA 6, 71-79, 1987

A:Title: Bovine and human cDNA sequences encoding a putative benzodiazepine receptor 11

A:Reference number: A26448; MUID:87161236; PMID:2881742

A:Accession: A26448

A:Molecule type: mRNA

A:Residues: 1-87 <MWB>

A:Cross-references: EMBL:M15886; NID:g162978; PIDN:AAA30495.1; PID:g162979

R:Jensen, M.S.; Hojrup, P.; Rasmussen, J.T.; Knudsen, J.

Biochem. J. 284, 809-812, 1992
A:Title: Purification and characterization of variants of acyl-CoA-binding protein in th
A:Reference number: S23127; MUID:92322005; PMID:1622397
A:Accession: S23127
A:Molecule type: protein
A:Residues: 2-87 <JEN>
A:Experimental source: liver
R:Margardt, H.; Todaro, G.J.; Shoyab, M.
J. Biol. Chem. 261, 9727-9731, 1986
A:Title: Complete amino acid sequences of bovine and human endozepines: homology with re
A:Reference number: A25832; MUID:86278003; PMID:3525533
A:Accession: A25832
A:Molecule type: protein
A:Residues: 2-87 <MAR>
A:Experimental source: brain
R:Mikkelsen, J.; Hojrup, P.; Nielsen, P.F.; Roepstorff, P.; Knudsen, J.
Biochem. J. 245, 857-861, 1987
A:Title: Amino acid sequence of acyl-CoA-binding protein from cow liver.
A:Reference number: A27886; MUID:88024142; PMID:3663196
A:Accession: A27886
A:Molecule type: protein
A:Residues: 2-87 <MIK>
A:Experimental source: liver
R:Besman, M.J.; Yanagibashi, K.; Lee, T.D.; Kawamura, M.; Hall, P.F.; Shively, J.E.
Proc. Natl. Acad. Sci. U.S.A. 86, 4897-4901, 1989
A:Title: Identification of des-(Gly-Ile)-endozepine as an effector of corticotropin-depe
receptor
A:Reference number: A32944; MUID:89296911; PMID:2544879
A:Accession: A32944
A:Molecule type: protein
A:Residues: 2-85 <BES>
A:Experimental source: adrenal cortex
C:Comment: This protein may take part in the modulation of GABA-ergic transmission in th
C:Superfamily: endozepine; endozepine homology
C:Keywords: acetylated amino end; acyl-CoA binding
F:2-87/Product: endozepine #status experimental <MAT>
F:2-85/Product: endozepine, short form #status experimental <MA2>
F:19-75/Domain: endozepine homology <ND2>
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

Query Match 1.6%; Score 8; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGRKAKMDA 315
DB 51 KGRKAKMDA 58

RESULT 8
A57711.
diazepam-binding inhibitor - laughing frog
C:Species: Rana ridibunda (laughing frog)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 13-Aug-1999
C:Accession: A57711
R:Uhrmann, I.; Plagevent, J.C.; Tostivint, H.; Rajmakers, R.; Tonon, M.C.; Conlon, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 6899-6903, 1994
A:Title: Frog diazepam-binding inhibitor: peptide sequence, cDNA cloning, and expression
A:Reference number: A57711; MUID:94316605; PMID:8041717
A:Accession: A57711
A:Molecule type: preliminary
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-88 <LIH>
A:Cross-references: GB:U09205; NID:g484097; PIDN:AAB60606.1; PID:g514280
C:Superfamily: endozepine; endozepine homology
F:20-76/Domain: endozepine homology <ND2>

Query Match 1.6%; Score 8; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGRKAKMDA 315
DB 51 KGRKAKMDA 58

DB 52 KGRKAKMDA 59

RESULT 9
C84710
hypothetical protein At2g30590 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84710
R:Lin, X.; Kal, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayem, L.; Tallon, L.;
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617979
A:Accession: C84710
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-380 <STO>
A:Cross-references: GB:AE002093; NID:g1946360; PIDN:AAB63078.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g30590
A:Map position: 2

Query Match 1.6%; Score 8; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SGFOELSL 28
DB 103 SGFOELSL 110

RESULT 10
S57589
probable membrane protein YMR221c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YMR959.03c
C:Species: Saccharomyces cerevisiae
C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 19-Apr-2002
C:Accession: S57589
R:Skellton, J.; Churcher, C.M.
submitted to the EMBL Data Library, June 1995
A:Reference number: S57587
A:Accession: S57589
A:Molecule type: DNA
A:Residues: 1-504 <SKF>
A:Cross-references: EMBL:Z49393; NID:g887599; PID:g887602; GSPDB:GN00013; MIPS:YMR221c
A:Experimental source: strain AB972
C:Genetics:
A:Gene: MIPS:YMR221c
A:Cross-references: SGD:S0004834
A:Map position: 13R
C:Superfamily: Saccharomyces cerevisiae probable membrane protein YMR221c
C:Keywords: transmembrane protein
F:12-28/Domain: transmembrane #status predicted <TM1>
F:65-81/Domain: transmembrane #status predicted <TM2>
F:93-109/Domain: transmembrane #status predicted <TM3>
F:127-143/Domain: transmembrane #status predicted <TM4>
F:190-206/Domain: transmembrane #status predicted <TM5>
F:306-322/Domain: transmembrane #status predicted <TM6>
F:349-365/Domain: transmembrane #status predicted <TM7>
F:374-390/Domain: transmembrane #status predicted <TM8>
F:399-415/Domain: transmembrane #status predicted <TM9>
F:467-483/Domain: transmembrane #status predicted <TM10>

Query Match 1.6%; Score 8; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 YLVGYTL 247
DB 121 YLVGYTL 128

RESULT 11

A43713
 C:Species: calcium-dependent protein kinase (EC 2.7.1.-) - soybean
 C:Date: 03-Mar-1993 #sequence_revision 14-Jul-1994 #text_change 11-Jun-1999
 C:Accession: A43713
 R:Harper, J.F.; Sussman, M.R.; Schaller, G.E.; Putnam-Evans, C.; Charbonneau, H.; Harmon Science 252, 951-954, 1991
 A:Title: A calcium-dependent protein kinase with a regulatory domain similar to calmodulin
 A:Reference number: A43713; MUID:91240279; PMID:1852075
 A:Accession: A43713
 A:Molecule type: mRNA
 A:Residues: 1-508 <HAR>
 A:Cross-references: EMBL:U64987; NID:g169930; PIDN:AA00806.1; PID:g169931
 C:Superfamily: calcium-dependent protein kinase; calmodulin repeat homology; protein kin
 C:Keywords: ATP; calcium binding; EF hand; phosphotransferase; serine/threonine-specific
 F:32-292/Domain: protein kinase homology <KIN>
 F:40-48/Region: protein kinase ATP-binding motif
 F:335-367/Domain: calmodulin repeat homology <EF1>
 F:371-403/Domain: calmodulin repeat homology <EF2>
 F:407-439/Domain: calmodulin repeat homology <EF3>
 F:441-473/Domain: calmodulin repeat homology <EF4>
 F:63/Active site: Lys #status predicted

Query Match 1.6%; Score 8; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 EROARLI 401
 |||||
 Db 132 EROARLI 139

RESULT 12

D95337
 C:Species: manganese transport protein Sma115 (imported) - Sinorhizobium meliloti (strain
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: D95337
 R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
 .; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yen, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9683-9688, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: D95337
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-547 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK65262.1; PID:g14523713; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galbert, F.; Finan, T.M.; Long, S.R.; Phibler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pola, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeLaure,
 hebaul, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma115
 A:Genome: plasmid

Query Match 1.6%; Score 8; DB 2; Length 547;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 LIGLAVDY 159
 |||||
 Db 472 LIGLAVDY 479

RESULT 13

H97854
 C:Species: hypothetical protein RC1240 (imported) - Rickettsia conorii (strain Malish 7)
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
 C:Accession: H97854
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
 Science 293, 2093-2098, 2001
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: H97854
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-54 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AAL03778.1; PID:g15620374; GSPDB:GN00173
 C:Genetics:
 A:Gene: RC1240

Query Match 1.4%; Score 7; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AEGDEGT 11
 |||||
 Db 25 AEGDEGT 31

RESULT 14

D95140
 C:Species: hypothetical protein SP120 (imported) - Streptococcus pneumoniae (strain TIGR4)
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: D95140
 R:Retelien, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: D95140
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-55 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75317.1; PID:g14972691; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP120

Query Match 1.4%; Score 7; DB 2; Length 55;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 LSGGLSL 367
 |||||
 Db 4 LSGGLSL 10

RESULT 15

S63592
 C:Species: acyl-coenzyme A-binding protein - dog
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
 C:Accession: S63592
 R:Kragelund, B.B.; Hojrup, P.; Skou Jensen, M.; Karlskov Schierling, C.; Juul, E.; Knud
 J. Mol. Biol. 256, 187-200, 1996
 A:Title: Fast and one-step folding of closely and distantly related homologous proteins
 A:Reference number: S63592; MUID:96190960; PMID:8609609
 A:Accession: S63592
 A:Molecule type: protein
 A:Residues: 1-86 <KGA>
 C:Superfamily: endozepine; endozepine homology
 C:Keywords: acyl-CoA binding

F:18-74/Domain: endozeptine homology <NDZ>

Query Match 1.4%; Score 7; DB 2; Length 86;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 309 GKAKMDA 315

|||||

Db 51 GKAKMDA 57

Search completed: November 21, 2003, 14:24:33
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 21, 2003, 14:17:14 ; Search time 17 Seconds
(without alignments)
1347.177 Million cell updates/sec

Title: US-09-938-803-6
Perfect score: 487
Sequence: 1 MASSABGDEGTVALAGVLQ.....ILPYIKAMVPRKELKME 487

Scoring table: OLIGO
Gapco 60.0, Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	1.8	556	1 NUMB_DROME	P16554 drosophila
2	8	1.6	86	1 ACBP_BOVIN	P07107 bos taurus
3	8	1.6	86	1 ACBP_CHAVI	P82934 chaetophrac
4	8	1.6	86	1 ACBP_CHICK	O9918 gallus gall
5	8	1.6	86	1 ACBP_PIG	P12026 sus scrofa
6	8	1.6	87	1 ACBP_RANR1	P45883 rana ridibu
7	8	1.6	283	1 H1S1_SULTO	O97023 sulfobus
8	8	1.6	310	1 SY81_ARATH	P59277 arabidopsis
9	8	1.6	380	1 WR21_ARATH	O04336 arabidopsis
10	8	1.6	504	1 YM68_YEAST	O04991 saccharomyc
11	8	1.6	508	1 CDPK_SOYBN	P28583 glycine max
12	7	1.4	64	1 AMP2_MELGA	P80392 melagris g
13	7	1.4	64	1 GLI3_CHICK	P46158 gallus gall
14	7	1.4	86	1 ACBP_CANFA	O9tcx6 canis fami
15	7	1.4	86	1 ACBP_HUMAN	P07108 homo sapien
16	7	1.4	86	1 ACBP_MOUSE	P31786 mus musculu
17	7	1.4	86	1 ACBP_RAT	P11030 rattus norv
18	7	1.4	107	1 YG1Y_YEAST	P53232 saccharomyc
19	7	1.4	126	1 C11X_HUMAN	O15332 homo sapien
20	7	1.4	134	1 NUDC_BUCAL	P57552 buchnera ap
21	7	1.4	228	1 YD31_HBLPJ	O92193 helicobacte
22	7	1.4	242	1 PHOU_PSEAE	O51547 pseudomonas
23	7	1.4	256	1 PHOU_PSEAE	O92410 pseudomonas
24	7	1.4	260	1 COAT_TYLCV	P27256 tomato yell
25	7	1.4	275	1 NADE_ECOLI	O8x229 escherichia
26	7	1.4	275	1 NADE_ECOLI	P18643 escherichia
27	7	1.4	275	1 NADE_SALTI	O82696 salmoneilla
28	7	1.4	294	1 CHIB_TOBAC	P29061 nicotiana t
29	7	1.4	294	1 CHIB_TOBAC	O92677 rickettsia
30	7	1.4	322	1 Y378_METUA	O57823 methanococc
31	7	1.4	323	1 C1TG_HAEDU	O30827 haemophilus
32	7	1.4	328	1 Z1PA_YERPE	P58492 yersinia pe
33	7	1.4	328	1 Z1PA_YERPE	P58492 yersinia pe

34	7	1.4	333	1 MOSA_RHIME	Q07607 rhizobium m
35	7	1.4	353	1 PRSW_BUCAP	O85913 buchnera ap
36	7	1.4	353	1 SOHB_HAEIN	P45315 haemophilus
37	7	1.4	354	1 CAHC_MOUSE	O8c185 mus musculu
38	7	1.4	355	1 CAHC_RABIT	O9m230 oryctolagus
39	7	1.4	369	1 DCHS_RHILLO	O98407 rhizobium l
40	7	1.4	376	1 H1S8_SULSO	O33770 sulfobus
41	7	1.4	397	1 CD61_SULSO	O98004 sulfobus
42	7	1.4	419	1 RRAA_SALTI	O84233 salmoneilla
43	7	1.4	419	1 RRAA_SALTI	P27031 salmoneilla
44	7	1.4	443	1 RRG2_HUMAN	P22932 homo sapien
45	7	1.4	447	1 RRG2_MOUSE	P20787 mus musculu

ALIGNMENTS

RESULT 1

ID	NUMB_DROME	STANDARD	PRT	556 AA
AC	P16554:			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Numb protein.			
GN	NUMB.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89324081; PubMed=2752427;			
RA	Uemura T., Shepherd S., Ackerman L., Jan L.Y., Jan Y.N.;			
RT	"Numb, a gene required in determination of cell fate during sensory			
RT	organ formation in Drosophila embryos.";			
RL	Cell 58:349-360(1989).			
RN	[2]			
RP	STRUCTURE BY NMR OF 64-210.			
RX	MEDLINE=99061335; PubMed=9846878;			
RA	Li S.-C., Zwahlen C., Vincent S.J., McGlade C.J., Kay L.E., Pawson T.,			
RT	Forman-Kay J.D.;			
RT	"Structure of a Numb protein domain-peptide complex suggests a basis for			
RT	diverse binding specificity.";			
RL	Nat. Struct. Biol. 5:1075-1083(1998).			
CC	-1- FUNCTION: NUMB IS REQUIRED IN DETERMINATION OF CELL FATE DURING			
CC	SENSORY ORGAN FORMATION IN DROSOPHILA EMBRYOS. IT FUNCTIONS IN			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative initiation;			
CC	Comment=2 isoforms, long/zygotic (shown here) and			
CC	Short/Maternal, may be produced by alternative initiation;			
CC	-1- SIMILARITY: Contains 1 PID domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-bb.ch).			
CC	-----			
DR	EMBL; M27815; AAA28730.1; .			
DR	PIR; A32466; A32466.			
DR	PDB; 1NMB; 04-NOV-98.			
DR	PDB; 1DDM; 12-APR-00.			
DR	flyBase; FBgn0002973; numb.			
DR	GO; GO:0045180; C:basal cortex; NAS.			
DR	InterPro; IPR006020; PTB_PID.			
DR	Pfam; PF00640; PID; 1.			
DR	SMART; SM00462; PTB; 1.			

DR PROSITE; PS01179; PID. 1.
 KW Developmental protein; Nuclear protein; ATP-binding;
 Alternative initiation; 3D-structure.
 FT CHAIN 1 556 NUMB PROTEIN, ISOFORM LONG.
 FT INIT MET 42 556 NUMB PROTEIN, ISOFORM SHORT.
 FT NP BIND 42 42 FOR ISOFORM SHORT.
 FT DOMAIN 22 29 ATP (POTENTIAL).
 FT DOMAIN 25 57 ARG/LYS-RICH (BASIC).
 FT DOMAIN 81 208 PID.
 FT TURN 66 68
 FT TURN 69 78
 FT HELIX 83 92
 FT STRAND 100 111
 FT TURN 112 113
 FT STRAND 119 122
 FT STRAND 129 132
 FT STRAND 138 138
 FT TURN 144 146
 FT STRAND 163 165
 FT TURN 167 168
 FT STRAND 172 179
 FT HELIX 188 194
 FT TURN 195 198
 FT TURN 200 201
 SQ SEQUENCE 556 AA; 60628 MW; 4FECAB9C98FE71 CRC64;
 Query Match 1.8%; Score 9; DB 1; Length 556;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 360 ELISGSLSL 368
 DB 369 ELISGSLSL 377

RESULT 2
 ACPB BOVIN STANDARD; PRT; 86 AA.
 ID ACPB BOVIN STANDARD; PRT; 86 AA.
 AC P07107;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Acyl-CoA-binding protein (ACBP) (Diazepam binding inhibitor) (DBI)
 DE (Endozepine) (EP).
 GN DBI.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87161236; PubMed=2881742;
 RA Webb N.R., Rose T.W., Malik N., Margardt H., Shoyab M., Todaro G.J.,
 RA Lee D.C.;
 RA "Bovine and human cDNA sequences encoding a putative benzodiazepine
 RT receptor ligand."
 RT DNA 6:71-79(1987).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=86278003; PubMed=3525533;
 RA Margardt H., Todaro G.J., Shoyab M.;
 RA "Complete amino acid sequences of bovine and human endozepines.
 RT Homology with rat diazepam binding inhibitor."
 RT J. Biol. Chem. 261:9727-9731(1986).
 RN [3]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=88024142; PubMed=3663196;
 RA Mikkelsen U., Hojrup P., Nielsen P.F., Roepstorff P., Knudsen U.;
 RA "Amino acid sequence of acyl-CoA-binding protein from cow liver."
 RT Biochem. J. 245:857-861(1987).

RN [4]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=92322005; PubMed=1622397;
 RA Jensen M.S., Hojrup P., Raemussen U.T., Knudsen U.;
 RA "Purification and characterization of variants of acyl-CoA-binding
 RT protein in the bovine liver."
 RL Biochem. J. 284:809-812(1992).
 RN [5]
 RP STRUCTURE BY NMR.
 RX MEDLINE=92031512; PubMed=1931985;
 RA Andersen K.V., Ludvigsen S., Mandrup S., Knudsen U., Poulsen F.M.;
 RA "The secondary structure in solution of acyl-coenzyme A binding
 RT protein from bovine liver using 1H nuclear magnetic resonance
 RL spectroscopy."
 RL Biochemistry 30:10654-10663(1991).
 RN [6]
 RP STRUCTURE BY NMR.
 RX MEDLINE=92389322; PubMed=1518047;
 RA Andersen K.V., Poulsen F.M.;
 RA "Three-dimensional structure in solution of acyl-coenzyme A binding
 RT protein from bovine liver."
 RL J. Mol. Biol. 226:1131-1141(1992).
 RN [7]
 RP STRUCTURE BY NMR.
 RX MEDLINE=93364192; PubMed=8358232;
 RA Andersen K.V., Poulsen F.M.;
 RA "The three-dimensional structure of acyl-coenzyme A binding protein
 RT from bovine liver: structural refinement using heteronuclear
 RL multidimensional NMR spectroscopy."
 RL J. Biomol. NMR 3:271-284(1993).
 CC - FUNCTION: BINDS MEDIUM- AND LONG-CHAIN ACYL-COA ESTERS WITH VERY
 CC HIGH AFFINITY AND MAY FUNCTION AS AN INTRACELLULAR CARRIER OF
 CC BENZODIAZEPINE (BZD) RECOGNITION SITE LOCATED ON THE GABA TYPE A
 CC RECEPTOR. IT IS THEREFORE POSSIBLE THAT THIS PROTEIN ALSO ACTS
 CC AS A NEUROPEPTIDE TO MODULATE THE ACTION OF THE GABA RECEPTOR.
 CC - SUBUNIT: Monomer.
 CC - SIMILARITY: Belongs to the ACPB family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M15886; AAA30495.1; -
 CC PIR; A26448; NZBO.
 DR PDB; 2ABD; 15-JUL-93.
 DR PDB; 1ACA; 30-APR-94.
 DR PDB; 1HB8; 11-MAR-02.
 DR InterPro; IPR000582; Ac_coA_bind_prot.
 DR Pfam; PF00887; ACPB; 1.
 DR PRINTS; PR00689; ACOABINDINGP.
 DR ProDom; PD351532; Ac_coA_bind_prot; 1.
 DR PROSITE; PS00880; ACPB; 1.
 KW PROSITE; PS00880; ACPB; 1.
 KW Transport; Lipid-binding; Acetylation; 3D-structure.
 FT INIT MET 0 0
 FT MOD RES 1 1 ACETYLATION.
 FT HELIX 4 14
 FT HELIX 21 35
 FT TURN 36 36
 FT TURN 48 61
 FT HELIX 62 63
 FT TURN 66 83
 FT HELIX 84 86
 FT TURN 84 86
 SQ SEQUENCE 86 AA; 9913 MW; 6B46B7C6C8CDE17 CRC64;
 Query Match 1.6%; Score 8; DB 1; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGKAKWDA 315
 |||||
 50 KGKAKWDA 57

Db

RESULT 3

ACBP_CHAVI STANDARD; PRT; 86 AA.

AC P82934;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acyl-CoA-binding protein (ACBP) (EP).
 GN DBI

OS Chætophractus villosus (South American armadillo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Edentata; Daelypodidae; Chætophractus.
 OX NCBI_TaxID=23080;
 RN [1]
 RP SEQUENCE, AND ACETYLATION.
 RC TISSUE=Harderian gland;
 RX MEDLINE=21240303; PubMed=11342056;
 RA Cavagnari B.M., Sterin-Speziale N., Affanni J.M., Knudsen J.,
 Santome J.A.;
 RT "Acyl-CoA-binding protein in the armadillo Harderian gland: its
 RT primary structure and possible role in lipid secretion."
 RL Biochim. Biophys. Acta 1545:314-325(2001).
 CC -1- FUNCTION: BINDS MEDIUM- AND LONG-CHAIN ACYL-COA ESTERS WITH VERY
 CC HIGH AFFINITY AND MAY FUNCTION AS AN INTRACELLULAR CARRIER OF
 CC ACYL-COA ESTERS.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SIMILARITY: Belongs to the ACBP family.
 DR HSSP: P07107; IACA.
 DR InterPro: IPR000582; Ac_coa_bind_prot.
 DR Pfam: PF00887; ACBP; 1.
 DR PRINTS: PR00689; ACOABINDINGP.
 DR PRODOM: PD351532; Ac_coa_bind_prot; 1.
 DR PROSITE: PS00880; ACBP; 1.
 KM Transprot; Lipid-binding; Acetylation.
 FT MOD_RES 1 1 ACETYLATION.
 SQ SEQUENCE 86 AA; 9891 MW; 6D6B803B10A2C8CD CRC64;

QY 308 KGKAKWDA 315
 |||||
 50 KGKAKWDA 57

Db

RESULT 4

ACBP_CHICK STANDARD; PRT; 86 AA.

AC Q9PRL8;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acyl-CoA-binding protein (ACBP).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=9619060; PubMed=8609609;
 RA Kregelund B.B., Hoejrup P., Jensen M.S., Schjerling C.K., Juul E.,
 Knudsen J., Poulsen F.M.;
 RT "Fast and one-step folding of closely and distantly related homologous
 RT proteins of a four-helix bundle family."
 RL J. Mol. Biol. 256:187-200(1996).

CC -1- FUNCTION: BINDS MEDIUM- AND LONG-CHAIN ACYL-COA ESTERS WITH VERY
 CC HIGH AFFINITY AND MAY FUNCTION AS AN INTRACELLULAR CARRIER OF
 CC ACYL-COA ESTERS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the ACBP family.
 DR PIR: S63595; S63595.
 DR HSSP: P07107; IACA.
 DR InterPro: IPR000582; Ac_coa_bind_prot.
 DR Pfam: PF00887; ACBP; 1.
 DR PRINTS: PR00689; ACOABINDINGP.
 DR PRODOM: PD351532; Ac_coa_bind_prot; 1.
 DR PROSITE: PS00880; ACBP; 1.
 KM Transprot; Lipid-binding.
 SQ SEQUENCE 86 AA; 9645 MW; 5AC1824914DCEB3 CRC64;

QY 308 KGKAKWDA 315
 |||||
 50 KGKAKWDA 57

Db

RESULT 5

ACBP_PIG STANDARD; PRT; 86 AA.

AC P12026;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acyl-CoA-binding protein (ACBP) (Diazepam binding inhibitor) (DBI)
 DE (Endosepine) (EP) (Contains: DBI (32-86)).
 GN DBI.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Intestine;
 RX MEDLINE=88254787; PubMed=3289918;
 RA Chen Z.W., Agerberth B., Gell K., Andersson M., Mutt V.,
 Oestenson C.G., Efendic S., Barros-Soderling J., Persson B.,
 Joernvall H.;
 RT "Isolation and characterization of porcine diazepam-binding
 RT inhibitor, a polypeptide not only of cerebral occurrence but also
 RT common in intestinal tissues and with effects on regulation of
 RT insulin release."
 RL Eur. J. Biochem. 174:239-245(1988).
 RN [2]
 RP SEQUENCE OF 32-86.
 RC TISSUE=Intestine;
 RX MEDLINE=93387315; PubMed=8375398;
 RA Agerberth B., Boman A., Andersson M., Joernvall H., Mutt V.,
 Boman H.G.;
 RT "Isolation of three antibacterial peptides from pig intestine:
 RT gastric inhibitory polypeptide (7-42), diazepam-binding inhibitor
 RT (32-86) and a novel factor, peptide 3910."
 RL Eur. J. Biochem. 216:623-629(1993).
 CC -1- FUNCTION: BINDS MEDIUM- AND LONG-CHAIN ACYL-COA ESTERS WITH VERY
 CC HIGH AFFINITY AND MAY FUNCTION AS AN INTRACELLULAR CARRIER OF
 CC ACYL-COA ESTERS. IT IS ALSO ABLE TO DISPLACE DIAZEPAM FROM THE
 CC BENZODIAZEPINE (BZD) RECOGNITION SITE LOCATED ON THE GABA TYPE A
 CC RECEPTOR. IT IS THEREFORE POSSIBLE THAT THIS PROTEIN ALSO ACTS
 CC AS A NEUROPEPTIDE TO MODULATE THE ACTION OF THE GABA RECEPTOR.
 CC -1- FUNCTION: DBI (32-86) HAS ANTIBACTERIAL PROPERTIES.
 CC -1- SUBUNIT: Monomer.
 CC -1- SIMILARITY: Belongs to the ACBP family.
 DR PIR: S00805; NZSC.
 DR HSSP: P07107; IACA.
 DR InterPro: IPR000582; Ac_coa_bind_prot.
 DR Pfam: PF00887; ACBP; 1.
 DR PRINTS: PR00689; ACOABINDINGP.

```

DR Prodom; PD351532; Ac coa_bind_prot; 1.
DR PROSITE; PS00880; ACBP; 1.
KM Transport; Lipid-binding; Acetylation; Antibiotic.
FT PEPTIDE 32 86 DBI(32-86).
SQ SEQUENCE 86 AA; 9765 MW; EBF128771E07A0F7 CRC64;
Query Match 1.6%; Score 8; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGGAKWDA 315
DB 50 KGGAKWDA 57

RESULT 6
ACBP_RANRI STANDARD; PRT; 87 AA.
AC P45883;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Acyl-CoA-binding protein homolog (ACBP) (Diazepam binding inhibitor
DE homolog) (DBI).
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_Taxid=8406;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-39 AND 58-87.
RC TISSUE=Brain;
RX MEDLINE=94316605; PubMed=8041717;
RA Lihmann I., Plaquevent J.-C., Tostivint H., Rajmakers R.,
RA Tonon M.-C., Conlon J.M., Vaudry H.,
RA "Frog diazepam-binding inhibitor: peptide sequence, cDNA cloning, and
RA expression in the brain."
RL Proc. Natl. Acad. Sci. U.S.A. 91:6899-6903 (1994).
RL [2]
RP CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=92319720; PubMed=1341880;
RA Malagon M., Vaudry H., Vallarino M., Gracia-Navarro F., Tonon M.C.;
RA "Distribution and characterization of endopeptin-like
RA immunoreactivity in the central nervous system of the frog Rana
RA ridibunda."
RL Peptides 13:99-107 (1992).
RL [3]
CC -1- FUNCTION: MAY PLAY IMPORTANT FUNCTIONS IN THE CONTROL OF BRAIN AND
CC PITUITARY ACTIVITIES. MAY REGULATE GABA NEUROTRANSMISSION THROUGH
CC A PARACRINE AND/OR AUTOCRINE MECHANISM. MAY NOT BIND ACYL-COA
CC ESTERS.
CC -1- TISSUE SPECIFICITY: BRAIN. IS SELECTIVELY EXPRESSED IN GLIAL
CC CELLS.
CC -1- SIMILARITY: BELONGS TO THE ACBP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U09205; AAB60606.1; -
DR PIR; A57711; A57711.
DR HSSP; P07107; IACA.
DR InterPro; IPR000582; Ac_coa_bind_prot.
DR Pfam; PF00887; ACBP; 1.
DR PRODOM; PD00689; ACOABINDINGP.
DR Prodom; PD351532; Ac_coa_bind_prot; 1.
DR PROSITE; PS00880; ACBP; 1.
DR Transport; Lipid-binding.
FT INIT_MET 0

```

```

FT PEPTIDE 18 51 TRIKONTATEPEPTIDE (TTN)
FT PEPTIDE 34 51 (POTENTIAL).
SQ SEQUENCE 87 AA; 9677 MW; P6621EFC412054BA CRC64;
Query Match 1.6%; Score 8; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGGAKWDA 315
DB 51 KGGAKWDA 58

RESULT 7
HIS1_SULTO STANDARD; PRT; 283 AA.
AC Q97023;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP phosphoribosyltransferase (EC 2.4.2.17).
GN HISG OR S71459.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_Taxid=119955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / ?;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankel A., Kosugi H., Hoshoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudon Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermocidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140 (2001).
RL [2]
CC -1- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribose)-ATP + diphosphate =
CC ATP + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: Histidine biosynthesis; first step. Very important in the
CC regulation of histidine metabolism.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE ATP PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC LONG SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR HAMAP; MF_00079; -; 1.
DR InterPro; IPR001348; ATP_pnspho_trans.
DR Pfam; PF01634; HISG; 1.
DR Prodom; PD003516; ATP_pnspho_trans; 1.
DR TIGRFAMs; TIGR00070; HISG; 1.
DR PROSITE; PS01316; ATP_P_PHORIBOSYLTR; 1.
DR Histidine biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 283 AA; 30985 MW; 19C98FC028F73C8C CRC64;

Query Match 1.6%; Score 8; DB 1; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 SPTLSKLA 411

```


Db 235 SPTLSKLA 242

RESULT 8:
SYBL_ARATH STANDARD; PRT; 310 AA.

ID SYBL_ARATH Q9C812;
AC P59277; Q9C812;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Synlaxin 81 (Atsyp81).
GN Synlaxin 81 (Atsyp81). F19C24.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
White O., Alonso J., Altschul H., Araujo R., Bowman C.L., Brooks S.Y.,
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
Kim C.J., Koo H.L., Kremenetskaia I., Kutz D.B., Kwan A., Lam B.,
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maltz R., Marshall A.,
Miltchev J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
Uteichberg T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
"Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana";
RL Nature 408:816-820(2000).
[2]
SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
Feldmann K.;
"Full-length cDNA from Arabidopsis thaliana";
RT Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
RL FUNCTION: Vesicle trafficking protein that functions in the
secretory pathway (By similarity).
CC -1- SUBUNIT: Part of the t-SNARE complex (By similarity).
CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL AC025294; AAG50888.1; ALT_SEQ.
DR EMBL AY086559; AAM63632.1;
DR PROSITE PS00914; SYNTAXIN; FALSE_NEG.
KW Transport; Protein transport; Transmembrane; Coiled coil;
KW MultiGene family.
FT DOMAIN 1 289 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 290 310 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
(POTENTIAL).
FT DOMAIN 77 114 COILED COIL (POTENTIAL).
SQ SEQUENCE 310 AA; 35587 MW; 992C2B5F5B87B60 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 310;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 LGNKELSQ 363
Db 273 LGNKELSQ 280

RESULT 9
WR21_ARATH STANDARD; PRT; 380 AA.

ID WR21_ARATH O04336;
AC O04336;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable WRKY transcription factor 21 (WRKY DNA-binding protein 21).
GN WRKY21 OR ATG30590 OR T6B20.6 OR T6B20.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A. AND INDUCTION.
RC STRAIN=cv. Columbia;
RX MEDLINE=21342502; PubMed=11449049;
RA Yu D., Chen C., Chen Z.;
"Evidence for an important role of WRKY DNA binding proteins in the
regulation of NPR1 gene expression";
RL Plant Cell 13:1527-1540(2001).
[2]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fujii C.Y., Mason T.M., Bowman C.L., Barneshead M.E., Feldblyum T.V.,
Buell C.R., Ketchum K.A., Lee J.D., Ronning C.M., Koo H.L.,
Moffatt K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umeyam L.,
Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
Venter J.C.;
"Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RL Nature 402:761-768(1999).
[3]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
"Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
SSP consortium (SAIL/Stanford/PGEC).";
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Transcription factor. Interacts specifically with the W
box (5'-(T)TGAC(C/T)-3'), a frequently occurring elicitor-
responsive cis-acting element (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- INDUCTION: By salicylic acid.
CC -1- SIMILARITY: Belongs to the WRKY group II-d family.
CC -1- SIMILARITY: Contains 1 WRKY domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL AF272747; AAK28441.1;
DR EMBL U93215; AAB63078.1;
DR EMBL BT000231; AAN15550.1;
DR EMBL AY136441; AAM97106.1;
DR PIR; C84710; C84710.

DR InterPro: IPR003657; WRKY.
DR Pfam: PF03106; WRKY; 1.
DR PROSITE: PSS0811; WRKY; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 307 373 WRKY.
FT DOMAIN 154 178 POLY-GLN.
FT DOMAIN 190 195 POLY-GLN.
SQ SEQUENCE 380 AA; 43005 MW; AEE49BFD6EFD797 CRC64;
Query Match
Best Local Similarity 1.6%; Score 8; DB 1; Length 380;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 SSGFQELSL 28
DB 103 SSGFQELSL 110
RESULT 10
YMF8 YEAST STANDARD; PRT; 504 AA.
AC 004951;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothesized 56.2 kDa protein in ERG8-UBP8 intergenic region.
GN YMR221C OR YMR959.03C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagers K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII".
RL Nature 387:90-93(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z49939; CAA00192.1; -.
DR PIR: S57589; S57589.
DR SGD: S0004834; YMR221C.
KW Hypothesized protein; Transmembrane.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 92 112 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 187 207 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 345 365 POTENTIAL.
FT TRANSMEM 386 406 POTENTIAL.
FT TRANSMEM 422 442 POTENTIAL.
FT TRANSMEM 463 483 POTENTIAL.
SQ SEQUENCE 504 AA; 56177 MW; 38C5E75DFB874767 CRC64;
Query Match
Best Local Similarity 1.6%; Score 8; DB 1; Length 504;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 YLVGYTL 247
DB 121 YLVGYTL 128
RESULT 11
ID CDPK SOYBN STANDARD; PRT; 508 AA.
AC P28583;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-dependent protein kinase SK5 (EC 2.7.1.-) (CDPK).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxId=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RX MEDLINE=91240279; PubMed=1852075;
RA Harper J.F., Sussman M.R., Schaller G.E., Putnam-Evans C.,
RA Chabouneau H., Harmon A.C.;
RT "A calcium-dependent protein kinase with a regulatory domain similar
RT to calmodulin".
RL Science 252:951-954(1991).
CC -1- FUNCTION: MAY PLAY A ROLE IN SIGNAL TRANSDUCTION PATHWAYS THAT
CC INVOLVE CALCIUM AS A SECOND MESSENGER.
CC -1- ENZYME REGULATION: ACTIVATED BY CALCIUM. AUTOPHOSPHORYLATION MAY
CC PLAY AN IMPORTANT ROLE IN THE REGULATION OF THE KINASE ACTIVITY.
CC -1- TISSUE SPECIFICITY: FOUND THROUGHOUT THE PLANT.
CC -1- MISCELLANEOUS: THERE ARE MULTIPLE CDPK ISOFORMS IN SOYBEAN.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CANK SUBFAMILY.
CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M64987; AAB0806.1; -.
DR PIR: A43713; A43713.
DR HSSP: Q63450; 1A06.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR Pfam: PF00036; ehand; 4.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000012; EF-hand; 2.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00054; Eph; 4.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PSS00018; EF_HAND; 4.
DR PROSITE: PSS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PSS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PSS00111; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Calcium-binding; Phosphorylation.
FT DOMAIN 34 292 PROTEIN KINASE.
FT NP_BIND 40 48 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 158 158 BY SIMILARITY.
FT CA_BIND 348 359 EF-HAND 1 (POTENTIAL).
FT CA_BIND 384 395 EF-HAND 2 (POTENTIAL).
FT CA_BIND 420 431 EF-HAND 3 (POTENTIAL).
FT CA_BIND 454 465 EF-HAND 4 (POTENTIAL).
SQ SEQUENCE 508 AA; 57169 MW; AFD5C51224192E4 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 394 EROAARLI 401
 DB 132 EROAARLI 139

RESULT 12

AMP2_MELGA
 ID AMP2_MELGA STANDARD; PRT; 64 AA.
 AC P80392; O93506;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 OS Antimicrobial peptide THP2 precursor (Turkey heterophil peptide 2).
 OC Melasaris gallipavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
 NCBI_TaxID=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=98418188; PubMed=9745666;
 RA Brockus C.W., Harmon B.G., Jackwood M.W.;
 RT "Characterization of beta-defensin prepropeptide mRNA from chicken
 and turkey bone marrow";
 RL Anim. Genet. 29:283-289(1998).
 RN [2]
 RP SEQUENCE OF 29-48.
 RX MEDLINE=95053386; PubMed=7964174;
 RA Evans E.W., Beach G.G., Wunderlich J., Harmon B.G.;
 RT "Isolation of antimicrobial peptides from avian heterophils";
 RL J. Leukoc. Biol. 56:661-665(1994).
 CC -1- FUNCTION: BACTERICIDAL ACTIVITY; INHIBITS STAPHYLOCOCCUS AUREUS.
 CC -1- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF033338; AAC36054.1; -;
 KW Antibiotic; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 64 ANTIMICROBIAL PEPTIDE THP2.
 FT DISULFID 31 57 BY SIMILARITY.
 FT DISULFID 36 51 BY SIMILARITY.
 FT DISULFID 41 58 BY SIMILARITY.
 SQ SEQUENCE 64 AA; 7327 MW; 7E597FEE4FCA4995 CRC64;
 Query Match 1.4%; Score 7; DB 1; Length 64;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 366 SLFLAL 372
 DB 9 SLFLAL 15

RESULT 13

GLD3_CHICK
 ID GLD3_CHICK STANDARD; PRT; 64 AA.
 AC P46158;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 OS Gallinacin 2 precursor.
 DE Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=98418188; PubMed=9745666;
 RA Brockus C.W., Harmon B.G., Jackwood M.W.;
 RT "Characterization of beta-defensin prepropeptide mRNA from chicken
 and turkey bone marrow";
 RL Anim. Genet. 29:283-289(1998).
 RN [2]

RP SEQUENCE OF 29-64, AND CHARACTERIZATION.
 RC STRAIN=Cross Broiler-6; TISSUE=Leukocyte;
 RX MEDLINE=94200386; PubMed=8150085;
 RA Harwig S.S.L., Swiderek K.M., Kokryakov V.N., Tan L., Lee T.D.,
 RA Panyutich E.A., Aleshina G.M., Shamova O.V., Lehrer R.I.;
 RT "Gallinacins: cysteine-rich antimicrobial peptides of chicken
 leukocytes";
 RL FEBS Lett. 342:281-285(1994).
 CC -1- FUNCTION: HAS BACTERICIDAL ACTIVITY. POTENT AGAINST E. COLI ML-35,
 CC AND L. MONOCYTOGENES EGD BUT NOT C. ALBICANS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic granules.
 CC -1- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF033336; AAC36052.1; -;
 KW Antibiotic; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 64 GALLINACIN 2.
 FT DISULFID 31 57 BY SIMILARITY.
 FT DISULFID 36 51 BY SIMILARITY.
 FT DISULFID 41 58 BY SIMILARITY.
 SQ SEQUENCE 64 AA; 7141 MW; 73E147BE4FCD39E0 CRC64;

Query Match 1.4%; Score 7; DB 1; Length 64;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 366 SLFLAL 372
 DB 9 SLFLAL 15

RESULT 14

ACBP_CANFA
 ID ACBP_CANFA STANDARD; PRT; 86 AA.
 AC Q9T0X6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acyl-CoA-binding protein (ACBP) (Diazepam binding inhibitor) (DBI)
 DE (Endozepine) (EP).
 GN DBI.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96190960; PubMed=8609609;
 RA Krieglund B.B., Hoeyrup P., Jensen M.S., Schjerling C.K., Juul E.,
 RA Knudsen J., Poulsen F.M.;
 RT "Fast and one-step folding of closely and distantly related homologous
 proteins of a four-helix bundle family";

RL J. Mol. Biol. 256:187-200(1996).
 CC -1- FUNCTION: BINDS MEDIUM- AND LONG-CHAIN ACYL-COA ESTERS WITH VERY
 CC HIGH AFFINITY AND MAY FUNCTION AS AN INTRACELLULAR CARRIER OF
 CC ACYL-COA ESTERS.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SIMILARITY: Belongs to the ACPB family.
 DR PIR: S63592; S63592.
 DR HSSP: P07107; IACA.
 DR InterPro: IPR000582; Ac_coa_bind_prot.
 DR Pfam: PF00887; ACPB; 1.
 DR PRINTS: PR00689; ACOABINDINGP.
 DR ProDom: PD51532; Ac_coa_bind_prot; 1.
 DR PROSITE: PS00880; ACPB; 1.
 DR Transport; Lipid-binding; Acetylation.
 FT INIT MET 0
 FT MOD RES 1 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 86 AA; 9847 MW; 65465B207E9B59D CRC64;
 QY Query Match 1.4%; Score 7; DB 1; Length 86;
 Db Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 309 GSAKMDA 315
 51 GSAKMDA 57
 RESULT 15
 ACPB_HUMAN STANDARD; PRT; 86 AA.
 ID ACPB_HUMAN
 AC P07108; P08869;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Acyl-CoA-binding protein (ACBP) (Diazepam binding inhibitor) (DBI)
 DE (Benzodiazepine) (BP).
 GN DBI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE.
 RP TISSUE=Brain;
 RX MEDLINE=86278003; PubMed=3525533;
 RA Marguier H., Todaro G.J., Shoyab M.;
 RT "Complete amino acid sequences of bovine and human endopeptidases.
 RT Homology with rat diazepam binding inhibitor.";
 RT J. Biol. Chem. 261:9727-9731(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87161236; PubMed=2881742;
 RA Webb N.R., Rose T.M., Malik N., Marguier H., Shoyab M., Todaro G.J.,
 RA Lee D.C.;
 RT "Bovine and human cDNA sequences encoding a putative benzodiazepine
 RT receptor ligand.";
 RT DNA 6:71-79(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87016986; PubMed=3020548;
 RA Gray P.W., Glaister D., Seeburg P.H., Guidotti A., Costa E.;
 RT "Cloning and expression of cDNA for human diazepam binding inhibitor,
 RT a natural ligand of an allosteric regulatory site of the gamma-
 RT aminobutyric acid type A receptor.";
 RT Proc. Natl. Acad. Sci. U.S.A. 83:7547-7551(1986).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 RX MEDLINE=23388257; PubMed=12477932;
 RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buecaw K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Cantoni P., Prange C.,
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.O., Malek A.M., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik R.A.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalske U., Smalins D.E.,
 RA Scherich A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=95194304; PubMed=7534063;
 RA Kolner M., Rovio A., Alho H.;
 RT "The characterization of two diazepam binding inhibitor (DBI)
 RT transcripts in humans.";
 RL Biochem. J. 306:327-330(1995).
 CC -1- FUNCTION: BINDS MEDIUM- AND LONG-CHAIN ACYL-COA ESTERS WITH VERY
 CC HIGH AFFINITY AND MAY FUNCTION AS AN INTRACELLULAR CARRIER OF
 CC ACYL-COA ESTERS. IT IS ALSO ABLE TO DISPLACE DIAZEPAM FROM THE
 CC BENZODIAZEPINE (BZD) RECOGNITION SITE LOCATED ON THE GABA TYPE A
 CC RECEPTOR. IT IS THEREFORE POSSIBLE THAT THIS PROTEIN ALSO ACTS AS
 CC A NEUROPEPTIDE TO MODULATE THE ACTION OF THE GABA RECEPTOR.
 CC -1- SUBUNIT: Monomer.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Short;
 CC IsoId=P07108-1; Sequence=Displayed;
 CC Name=2; Synonyms=Long;
 CC IsoId=P07108-2; Sequence=VSP 000068;
 CC -1- SIMILARITY: Belongs to the ACPB family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@db-sib.ch).
 CC -----
 DR EMBL: M14200; AAAS2171.1; -;
 DR EMBL: M15887; AAA35788.1; -;
 DR EMBL: BC006466; AAH06466.1; -;
 DR PIR: B26448; NZHU.
 DR HSSP: P07107; IACA.
 DR Genew: HGNC:2690; DBI.
 DR MIM: 125950;
 DR GO: GO:0008025; F:diazepam-binding inhibitor activity; TAS.
 DR GO: GO:0005102; F:receptor binding activity; TAS.
 DR InterPro: IPR000582; Ac_coa_bind_prot.
 DR Pfam: PF00887; ACPB; 1.
 DR PRINTS: PR00689; ACOABINDINGP.
 DR ProDom: PD51532; Ac_coa_bind_prot; 1.
 DR PROSITE: PS00880; ACPB; 1.
 DR Transport; Lipid-binding; Acetylation; Alternative splicing.
 FT INIT MET 0
 FT MOD RES 1 1 ACETYLATION.
 FT VARSPPLIC 1 2 SO -> MMGDLMLPPASNPOTGTE (in isoform
 FT 2).
 SQ SEQUENCE 86 AA; 9913 MW; 6A973E406CB1D0E7 CRC64;
 QY Query Match 1.4%; Score 7; DB 1; Length 86;
 Db Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 309 GSAKMDA 315

Db |||
 51 GKAKMDA 57

Search completed: November 21, 2003, 14:23:07
Job time : 19 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:21:44 ; Search time 21 Seconds

(without alignments)
981.209 Million cell updates/sec

Title: US-09-938-803-6

Perfect score: 487

Sequence: 1 MASSAEDEGVVALAGVLQ.....ILPYIKAMVPRKELKKME 487

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	1.6	86	1 US-08-700-626-5	Sequence 5, Appl1
2	8	1.6	463	1 US-07-951-715A-25	Sequence 25, Appl1
3	8	1.6	463	2 US-08-459-448A-25	Sequence 25, Appl1
4	8	1.6	463	3 US-08-459-595A-25	Sequence 25, Appl1
5	8	1.6	463	3 US-08-459-504B-25	Sequence 25, Appl1
6	8	1.6	463	3 US-08-459-444-25	Sequence 0, Appl1
7	8	1.6	463	4 US-09-547-422-25	Sequence 73, Appl1
8	8	1.6	947	4 US-09-228-986-73	Sequence 49, Appl1
9	7	1.4	28	1 US-08-182-175A-4	Sequence 49, Appl1
10	7	1.4	28	1 US-08-182-175A-49	Sequence 83, Appl1
11	7	1.4	28	1 US-08-182-175A-83	Sequence 58, Appl1
12	7	1.4	28	1 US-08-474-633A-58	Sequence 58, Appl1
13	7	1.4	28	1 US-08-474-633A-73	Sequence 58, Appl1
14	7	1.4	28	4 US-08-823-771-58	Sequence 73, Appl1
15	7	1.4	28	4 US-08-823-771-73	Sequence 34, Appl1
16	7	1.4	28	4 US-09-351-657A-34	Sequence 35, Appl1
17	7	1.4	28	4 US-09-351-657A-35	Sequence 4, Appl1
18	7	1.4	28	5 PCT-US92-06412-4	Sequence 49, Appl1
19	7	1.4	28	5 PCT-US92-06412-49	Sequence 83, Appl1
20	7	1.4	28	5 PCT-US92-06412-83	Sequence 83, Appl1
21	7	1.4	31	1 US-08-190-802A-219	Sequence 219, App
22	7	1.4	31	3 US-08-477-346-219	Sequence 219, App
23	7	1.4	31	4 US-08-477-346-219	Sequence 219, App
24	7	1.4	31	4 US-08-487-072A-219	Sequence 10, Appl1
25	7	1.4	40	1 US-08-144-121-10	Sequence 10, Appl1
26	7	1.4	40	2 US-08-735-893-10	Sequence 403, App
27	7	1.4	60	4 US-09-149-476-403	

28	7	1.4	64	4 US-09-351-657A-4	Sequence 4, Appl1
29	7	1.4	64	4 US-09-351-657A-8	Sequence 8, Appl1
30	7	1.4	77	1 US-08-182-175A-57	Sequence 75, Appl1
31	7	1.4	77	1 US-08-474-633A-75	Sequence 57, Appl1
32	7	1.4	77	4 US-08-823-771-75	Sequence 57, Appl1
33	7	1.4	77	5 PCT-US92-06412-57	Sequence 4, Appl1
34	7	1.4	86	1 US-08-700-626-4	Sequence 55, Appl1
35	7	1.4	86	4 US-09-300-008B-55	Sequence 3, Appl1
36	7	1.4	104	1 US-08-700-626-3	Sequence 29325, A
37	7	1.4	187	4 US-09-252-991A-29325	Sequence 20, Appl1
38	7	1.4	188	2 US-08-484-993B-20	Sequence 20, Appl1
39	7	1.4	188	2 US-08-484-158B-20	Sequence 20, Appl1
40	7	1.4	188	2 US-08-484-596A-20	Sequence 20, Appl1
41	7	1.4	188	2 US-08-480-150A-20	Sequence 20, Appl1
42	7	1.4	188	3 US-08-458-731-20	Sequence 20, Appl1
43	7	1.4	188	3 US-08-149-223A-20	Sequence 20, Appl1
44	7	1.4	203	4 US-09-134-001C-3145	Sequence 3145, Ap
45	7	1.4	257	4 US-09-252-991A-23705	Sequence 23705, A

ALIGNMENTS

```
RESULT 1
US-08-700-626-5
; Sequence 5, Application US/08700626
; Patent No. 5734038
;
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN DBI/ACBP-LIKE PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700, 626
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0115 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118275
;
; US-08-700-626-5
;
; Query Match 1.6%; Score 8; DB 1; Length 86;
; Best Local Similarity 100.0%; Pred. No. 3.6;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 50 KOKAKMDA 57

RESULT 2

US-07-951-715A-25
Sequence 25, Application US/07951715A
Patent No. 5625136

GENERAL INFORMATION:

APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/951,715A
FILING DATE: 25-SEP-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..463
OTHER INFORMATION: /note="protein sequence for
OTHER INFORMATION: soybean CDPK as shown in figure 34."
US-07-951-715A-25

Query Match 1.6%; Score 8; DB 1; Length 463;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 394 ERQARLI 401

Db 111 ERQARLI 118

RESULT 3

US-08-459-448A-25
Sequence 25, Application US/08459448A
Patent No. 5859336

GENERAL INFORMATION:

APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 5859336artis Corporation
STREET: Patent & Trademark Dept., 520 White Plains
STREET: Rd., POB 2005
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-9005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,448A
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40403
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8582
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..463
OTHER INFORMATION: /note="protein sequence for
OTHER INFORMATION: soybean CDPK as shown in figure 34."
US-08-459-448A-25

Query Match 1.6%; Score 8; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 EROARLI 401
DB 111 EROARLI 118

RESULT 4

US-08-459-595A-25
Sequence 25, Application US/08459595A
Patent No. 6018104

GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6018104artis Corporation
STREET: Patent & Trademark Dept., 520 White Plains
STREET: Rd., POB 2005
CITY: Tarrytown
STATE: New York

COUNTRY: USA
ZIP: 10591-9005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,595A

FILING DATE: 02-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/951,715

FILING DATE: 25-SEP-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/772,027

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Pace, Gary M.

REGISTRATION NUMBER: 40403

REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (919)541-8582

TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 463 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: Protein

LOCATION: 1..463
OTHER INFORMATION: /note= "protein sequence for
soybean CDPK as shown in Figure 34."

Query Match 1.6%; Score 8; DB 3; Length 463;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 EROARLI 401
DB 111 EROARLI 118

RESULT 5

US-08-459-504B-25
Sequence 25, Application US/08459504B
Patent No. 6075185

GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6075185artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC

COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,504B

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/459,595

FILING DATE: 02-JUN-1995

APPLICATION NUMBER: US 07/951,715

FILING DATE: 25-SEP-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/772,027

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (919)541-8587

TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 463 amino acids

TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..463
OTHER INFORMATION: /note= "protein sequence for
US-08-459-504B-25 soybean CDPK as shown in Figure 34."
Query Match 1.6%; Score 8; DB 3; Length 463;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 394 EROARLI 401
DB 111 EROARLI 118
RESULT 6
US-08-459-444-25
Sequence 0, Application US/08459444A
Patent No. 6121014
GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Lauais, Karen L.
TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
NUCLEIC ACID CODING SEQUENCE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,444A
FILING DATE: 02-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805/PI/CGC1577/CIP/DIV6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: /note= "protein sequence for
soybean CDPK as shown in Figure 34."
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..463
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-08-459-444-25
Query Match 1.6%; Score 8; DB 3; Length 463;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 394 EROARLI 401
DB 111 EROARLI 118
RESULT 7
US-09-547-422-25
Sequence 0, Application US/09547422
Patent No. 6320100
GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Lauais, Karen L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/547,422
FILING DATE: 11-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805H
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: /note= "protein sequence for
soybean CDPK as shown in Figure 34."
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..463
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-547-422-25

Query Match 1.6%; Score 8; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 394 EROARLI 401
Db 111 EROARLI 118

RESULT 8
US-09-228-986-73
Sequence 73, Application US/09228986
Patent No. 6359198
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Niels
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
FILE REFERENCE: 11000/1020
CURRENT APPLICATION NUMBER: US/09/228,986
CURRENT FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 130
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 73
LENGTH: 947
TYPE: PRT
ORGANISM: Pinus radiata
US-09-228-986-73

Query Match 1.6%; Score 8; DB 4; Length 947;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 462 FLSLFIL 469
Db 10 FLSLFIL 17

RESULT 9
US-08-182-175A-4
Sequence 4, Application US/08182175A
Patent No. 5559223
GENERAL INFORMATION:
APPLICANT: Saverio Carl Falco
APPLICANT: Sharon J. Keeler
APPLICANT: Janet A. Rice
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing F
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. du Pont de Nemours and Company
STREET: 1007 Market Street
CITY: Wilmington
STATE: Delaware
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh System, 6.0
SOFTWARE: Microsoft Word, 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/182,175A
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/743,006

FILING DATE: 9 August 1991
ATTORNEY/AGENT INFORMATION:
NAME: Linda Axamethy Floyd
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: BB-1031
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 992-4929
TELEFAX: (302) 892-7949
TELEX: 835420
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..28
OTHER INFORMATION: /label= name
US-08-182-175A-4
OTHER INFORMATION: /note= "(SSP 8) 4"

Query Match 1.4%; Score 7; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 EKLKME 487
Db 3 EKLKME 9

RESULT 10
US-08-182-175A-49
Sequence 49, Application US/08182175A
Patent No. 5559223
GENERAL INFORMATION:
APPLICANT: Saverio Carl Falco
APPLICANT: Sharon J. Keeler
APPLICANT: Janet A. Rice
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing P
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. du Pont de Nemours and Company
STREET: 1007 Market Street
CITY: Wilmington
STATE: Delaware
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh System, 6.0
SOFTWARE: Microsoft Word, 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/182,175A
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/743,006
FILING DATE: 9 August 1991
ATTORNEY/AGENT INFORMATION:
NAME: Linda Axamethy Floyd
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: BB-1031
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 992-4929
TELEFAX: (302) 892-7949
TELEX: 835420
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD VERSION 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,633A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: BARBARA C. SIEGEL
REGISTRATION NUMBER: 30,684
REFERENCE/DOCKET NUMBER: BB-1037-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4931
TELEFAX: 302-773-0164
TELEX: 835420
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-474-633A-73

Query Match 1.4%; Score 7; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 481 EKLKME 487
|||||
Db 3 EKLKME 9

RESULT 14
US-08-823-771-58
Sequence 58, Application US/08823771
Patent No. 6459019
GENERAL INFORMATION:
APPLICANT: E. I. DU PONT DE NEMOURS AND
COMPANY
TITLE OF INVENTION: CHIMERIC GENES AND
METHODS FOR INCREASING
INCREASING THE LYSINE
AND THREONINE CONTENT
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS
AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD VERSION 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,771
FILING DATE: 24-Mar-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/474,633
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BARBARA C. SIEGEL
REGISTRATION NUMBER: 30,684
REFERENCE/DOCKET NUMBER: BB-1037-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4931
TELEFAX: 302-773-0164
TELEX: 835420

INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-08-823-771-58

Query Match 1.4%; Score 7; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 481 EKLKME 487
|||||
Db 10 EKLKME 16

RESULT 15
US-08-823-771-73
Sequence 73, Application US/08823771
Patent No. 6459019
GENERAL INFORMATION:
APPLICANT: E. I. DU PONT DE NEMOURS AND
COMPANY
TITLE OF INVENTION: CHIMERIC GENES AND
METHODS FOR INCREASING
INCREASING THE LYSINE
AND THREONINE CONTENT
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS
AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD VERSION 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,771
FILING DATE: 24-Mar-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/474,633
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BARBARA C. SIEGEL
REGISTRATION NUMBER: 30,684
REFERENCE/DOCKET NUMBER: BB-1037-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4931
TELEFAX: 302-773-0164
TELEX: 835420
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-08-823-771-73

Query Match 1.4%; Score 7; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 481 EKLKME 487

|||||
Db 3 EXTRACT 9

Search completed: November 21, 2003, 14:25:05
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 14:24:05 ; Search time 36 Seconds

(without alignments)
2469.624 Million cell updates/sec

Title: US-09-938-803-6

Perfect score: 487
Sequence: 1 MASSAEGDEGTVALAGVLAQ.....ILPYIKAMVPRKKKKME 487Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	487	100.0	487	9 US-09-938-803-6	Sequence 6, Appl1
2	56	11.5	57	12 US-10-029-386-29742	Sequence 29742, A
3	42	8.6	58	9 US-09-864-761-33415	Sequence 33415, A
4	8	1.6	86	9 US-09-859-101-5	Sequence 1, Appl1
5	8	1.6	86	10 US-09-987-108-1	Sequence 3, Appl1
6	8	1.6	86	10 US-09-987-108-3	Sequence 7, Appl1
7	8	1.6	86	10 US-09-987-108-8	Sequence 8, Appl1
8	8	1.6	86	10 US-09-987-108-9	Sequence 9, Appl1
9	8	1.6	86	10 US-09-987-108-20	Sequence 20, Appl1
10	8	1.6	463	11 US-09-988-462-25	Sequence 25, Appl1
11	8	1.6	609	12 US-10-032-585-7294	Sequence 7294, Ap
12	8	1.6	947	15 US-10-101-464A-73	Sequence 73, Appl1
13	8	1.4	20	11 US-09-814-604-26	Sequence 26, Appl1
14	7	1.4	20	11 US-09-922-226-12	Sequence 12, Appl1
15	7	1.4	20	11 US-09-922-226-12	Sequence 12, Appl1

16	7	1.4	20	11	US-09-922-226-165	Sequence 165, App
17	7	1.4	28	10	US-09-982-172-226	Sequence 226, Appl
18	7	1.4	28	15	US-10-023-066A-58	Sequence 58, Appl
19	7	1.4	28	15	US-10-023-066A-73	Sequence 73, Appl
20	7	1.4	33	9	US-09-281-717-34	Sequence 34, Appl
21	7	1.4	35	9	US-09-864-761-37156	Sequence 37156, A
22	7	1.4	37	10	US-09-982-172-94	Sequence 94, Appl
23	7	1.4	45	15	US-10-074-475-243	Sequence 243, App
24	7	1.4	50	11	US-09-974-879-225	Sequence 225, App
25	7	1.4	51	11	US-09-105-736-225	Sequence 225, App
26	7	1.4	54	15	US-10-227-629-9	Sequence 9, Appl1
27	7	1.4	60	11	US-09-809-991-403	Sequence 403, App
28	7	1.4	60	12	US-09-882-171-403	Sequence 403, App
29	7	1.4	68	14	US-10-001-857-130	Sequence 130, App
30	7	1.4	77	15	US-10-023-066A-75	Sequence 75, Appl
31	7	1.4	78	9	US-09-864-761-37650	Sequence 37650, A
32	7	1.4	86	9	US-09-859-101-4	Sequence 4, Appl1
33	7	1.4	86	10	US-09-987-108-2	Sequence 2, Appl1
34	7	1.4	86	10	US-09-987-108-4	Sequence 4, Appl1
35	7	1.4	86	10	US-09-987-108-5	Sequence 5, Appl1
36	7	1.4	86	10	US-09-987-108-6	Sequence 6, Appl1
37	7	1.4	87	12	US-10-334-038-12	Sequence 12, Appl
38	7	1.4	104	9	US-09-859-101-3	Sequence 3, Appl1
39	7	1.4	104	15	US-10-136-807A-6	Sequence 6, Appl1
40	7	1.4	106	15	US-10-106-698-6734	Sequence 6734, Ap
41	7	1.4	111	9	US-09-925-299-1143	Sequence 1143, Ap
42	7	1.4	111	11	US-09-925-299-1143	Sequence 1143, Ap
43	7	1.4	158	12	US-10-032-585-7497	Sequence 7497, Ap
44	7	1.4	212	9	US-09-864-761-46654	Sequence 46654, A
45	7	1.4	275	9	US-09-815-242-10184	Sequence 10184, A

ALIGNMENTS

```
RESULT 1
US-09-938-803-6
; Sequence 6, Application US/09938803
; Patent No. US20020076762A1
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Reddy, Roopa
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Yang, Junning
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: FULL-LENGTH EXPRESSED GENETIC MARKERS
; FILE REFERENCE: PF-0695 US
; CURRENT APPLICATION NUMBER: US/09/938,803
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/311,894
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inocyte Clone 1928920
US-09-938-803-6

Query Match      100.0%; Score 487, DB 9, length 487,
Best Local Similarity 100.0%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MASSAEGDEGTVALAGVLSGFOELSLNKATSLGASGQARLRIISIFLGGPALFYRH 60
        |||
        1 MASSAEGDEGTVALAGVLSGFOELSLNKATSLGASGQARLRIISIFLGGPALFYRH 60
QY      61 YLFYKETYLIHLFHTFTGSLIAYFNGQLYHSLCLIVQLPILRLMGRITTAVALTTFCF 120
```

Db 61 YLFYETVYIHLPHFTGTGSIAYFNFQNGLYHSLSLCTIVQFLILRMGRTITAVLTTPRF 120
Qy 121 QMAVLLAGYATGATGNYDIDKMTMPHCVLTKLIGLAVDYPDGKQNSLSSQOKYAIRG 180
Db 121 QMAVLLAGYATGATGNYDIDKMTMPHCVLTKLIGLAVDYPDGKQNSLSSQOKYAIRG 180
Qy 181 VPSLLEVGAFSYFYGAFLVGPQFSNMHWKLVQGEILIDIPGKI PMSIIPALRSLGIFY 240
Db 181 VPSLLEVGAFSYFYGAFLVGPQFSNMHWKLVQGEILIDIPGKI PMSIIPALRSLGIFY 240
Qy 241 LVGYTLSPHITEDYLTEDYDNHHPFRCMVMILMGKFLYKYVTGMLVTEGVCILITGL 300
Db 241 LVGYTLSPHITEDYLTEDYDNHHPFRCMVMILMGKFLYKYVTGMLVTEGVCILITGL 300
Qy 301 GENGEEKGAQKADACANKKWLFEITNPRFTGTIASFNININAMAVARYFKRLKLGKKE 360
Db 301 GENGEEKGAQKADACANKKWLFEITNPRFTGTIASFNININAMAVARYFKRLKLGKKE 360
Qy 361 LSQGLSLFLALMHGHSGLVGCPOEFLIVVERQAALIQESPTLSKLAITVLOPRY 420
Db 361 LSQGLSLFLALMHGHSGLVGCPOEFLIVVERQAALIQESPTLSKLAITVLOPRY 420
Qy 421 YLVQOTIHLFMGYSMTAFCLFTWPKMLKYKSIYFLGHIFFLSLFLIPYIHKMVPK 480
Db 421 YLVQOTIHLFMGYSMTAFCLFTWPKMLKYKSIYFLGHIFFLSLFLIPYIHKMVPK 480
Qy 481 EKLKXME 487
Db 481 EKLKXME 487

RESULT 2

US-10-029-386-29742
; Sequence 29742, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34286
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29742
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR12.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.5
; OTHER INFORMATION: SWISSPROT HIT: P54336, EVALU 8.00e+00
US-10-029-386-29742

Query Match 11.5%; Score 56; DB 12; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.8e-46;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 292 EGVYCLITGIGFNGFEKGAQKADACANKKWLFEITNPRFTGTIASFNININAMAVAR 347
Db 2 EGVYCLITGIGFNGFEKGAQKADACANKKWLFEITNPRFTGTIASFNININAMAVAR 347

RESULT 3

US-09-864-761-33415
; Sequence 33415, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEWICA-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33415
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006512.12
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HELA100, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EST_HUMAN HIT: W26367.1, EVALU 3.00e-26
; OTHER INFORMATION: SWISSPROT HIT: P43288, EVALU 4.50e+00
US-09-864-761-33415

Query Match 8.6%; Score 42; DB 9; Length 58;

Best Local Similarity 100.0%; Pred. No. 9,7e-33;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 396 QARLQESPTLSKLAITVLOPEFYLVQQTTHMLFMCYSMT 437
Db 17 QARLQESPTLSKLAITVLOPEFYLVQQTTHMLFMCYSMT 58

RESULT 4

US-09-859-101-5
Sequence 5, Application US/09859101
Patent No. US2002068829A1
GENERAL INFORMATION:

APPLICANT: Au-Young, Janice
Hillman, Jennifer L.
Goli, Surya K.

TITLE OF INVENTION: NOVEL HUMAN DBI/ACBP-LIKE PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto

STATE: CA
COUNTRY: U.S.
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,101
FILING DATE: 14-May-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/700,626
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0115 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555

TELEFAX: 415-845-0555
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 86 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
IMMEDIATE SOURCE:

LIBRARY: Genbank
CLONE: 118275

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-859-101-5.

Query Match 1.6%; Score 8; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGKAKMDA 315
Db 50 KGKAKMDA 57

RESULT 5

US-09-987-108-1
Sequence 1, Application US/09987108
Patent No. US20020142347A1

GENERAL INFORMATION:

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: NEERGAARD, Thomas B.F.
TITLE OF INVENTION: BIOSENSOR
FILE REFERENCE: KNUDSENIA

CURRENT APPLICATION NUMBER: US/09/987,108
CURRENT FILING DATE: 2001-11-13

PRIOR APPLICATION NUMBER: US 60/262,366
PRIOR FILING DATE: 2001-01-19

PRIOR APPLICATION NUMBER: DK PA2000 01638
PRIOR FILING DATE: 2000-11-10

NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn version 3.1

SEQ ID NO 1
LENGTH: 86

TYPE: PRT
ORGANISM: Bos taurus

US-09-987-108-1

Query Match 1.6%; Score 8; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGKAKMDA 315
Db 50 KGKAKMDA 57

RESULT 6

US-09-987-108-3
Sequence 3, Application US/09987108
Patent No. US20020142347A1

GENERAL INFORMATION:

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

TITLE OF INVENTION: BIOSENSOR
FILE REFERENCE: KNUDSENIA

CURRENT APPLICATION NUMBER: US/09/987,108
CURRENT FILING DATE: 2001-11-13

PRIOR APPLICATION NUMBER: US 60/262,366
PRIOR FILING DATE: 2001-01-19

PRIOR APPLICATION NUMBER: DK PA2000 01638
PRIOR FILING DATE: 2000-11-10

NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn version 3.1

SEQ ID NO 3
LENGTH: 86

TYPE: PRT
ORGANISM: Sus scrofa

US-09-987-108-3

Query Match 1.6%; Score 8; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KGKAKMDA 315
Db 50 KGKAKMDA 57

US-09-987-108-7
Sequence 7, Application US/09987108
Patent No. US20020142347A1

GENERAL INFORMATION:

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

APPLICANT: Knudsen, Jens
Wadum, Maiken C.T.

TITLE OF INVENTION: BIOSENSOR
FILE REFERENCE: KNUDSENIA

CURRENT APPLICATION NUMBER: US/09/987,108
CURRENT FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/262,366
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: DK PA2000 01638
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Terrapene carolina
US-09-987-108-7

Query Match 1.6%; Score 8; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KKGAKWDA 315
DB 50 KKGAKWDA 57

RESULT 8
US-09-987-108-8
; Sequence 8, Application US/09987108
; Patent No. US20020142347A1
; GENERAL INFORMATION:
; APPLICANT: KNUDSEN, Jens
; APPLICANT: WADUM, Maiken C.T.
; APPLICANT: VILLADSEN, Jens
; APPLICANT: NEERGAARD, Thomas B.F.
; TITLE OF INVENTION: BIOSENSOR
; FILE REFERENCE: KNUDSENIA
; CURRENT APPLICATION NUMBER: US/09/987,108
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/262,366
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: DK PA2000 01638
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Anas platyrhynchos
US-09-987-108-8

Query Match 1.6%; Score 8; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KKGAKWDA 315
DB 50 KKGAKWDA 57

RESULT 9
US-09-987-108-9
; Sequence 9, Application US/09987108
; Patent No. US20020142347A1
; GENERAL INFORMATION:
; APPLICANT: KNUDSEN, Jens
; APPLICANT: WADUM, Maiken C.T.
; APPLICANT: VILLADSEN, Jens
; APPLICANT: NEERGAARD, Thomas B.F.
; TITLE OF INVENTION: BIOSENSOR
; FILE REFERENCE: KNUDSENIA
; CURRENT APPLICATION NUMBER: US/09/987,108
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/262,366
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: DK PA2000 01638
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-987-108-9

Query Match 1.6%; Score 8; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KKGAKWDA 315
DB 50 KKGAKWDA 57

RESULT 10
US-09-987-108-20
; Sequence 20, Application US/09987108
; Patent No. US20020142347A1
; GENERAL INFORMATION:
; APPLICANT: KNUDSEN, Jens
; APPLICANT: WADUM, Maiken C.T.
; APPLICANT: VILLADSEN, Jens
; APPLICANT: NEERGAARD, Thomas B.F.
; TITLE OF INVENTION: BIOSENSOR
; FILE REFERENCE: KNUDSENIA
; CURRENT APPLICATION NUMBER: US/09/987,108
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/262,366
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: DK PA2000 01638
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Rana sp.
US-09-987-108-20

Query Match 1.6%; Score 8; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 KKGAKWDA 315
DB 50 KKGAKWDA 57

RESULT 11
US-09-988-462-25
; Sequence 25, Application US/09988462
; Publication No. US20030046726A1
; GENERAL INFORMATION:
; APPLICANT: Kozziel, Michael G.
; Desai, Nalin M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launls, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Syngenta Biotechnology, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC

COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/988,462
FILING DATE: 20-NOV-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/547,422
FILING DATE: 11-APR-2000
APPLICATION NUMBER: US 08/459,504
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-188051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid.
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..463
OTHER INFORMATION: /notes "protein sequence for
soybean CDPK as shown in Figure 34."
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-988-462-25
Query Match 1.6%; Score 8; DB 11; Length 463;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 394 ERQARLI 401
Db 111 ERQARLI 118
RESULT 12
US-10-032-585-7294
Sequence 7294, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiansg
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: Patent version 3.1
SEQ ID NO 7294
LENGTH: 609
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7294

Query Match 1.6%; Score 8; DB 12; Length 609;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 296 ILTGIGFN 303
Db 208 ILTGIGFN 215
RESULT 13
US-10-101-464A-73
Sequence 73, Application US/10101464A
Publication No. US20030046728A1
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
and Their Use in the Modification of Plant Cell Signaling
FILE REFERENCE: 11000.1020C2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
PRIOR FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 73
LENGTH: 947
TYPE: PRT
ORGANISM: Pinus radiata
US-10-101-464A-73
Query Match 1.6%; Score 8; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 462 FLSLFIL 469
Db 10 FLSLFIL 17
RESULT 14
US-09-814-604-26
Sequence 26, Application US/09814604
Publication No. US20030003517A1
GENERAL INFORMATION:
APPLICANT: Klein, Elliott S.
APPLICANT: Chandraratna Roshantha A.
TITLE OF INVENTION: Methods of Detecting Dissociated Nuclear
Hormone Receptor Ligands
FILE REFERENCE: P-AR 4528
CURRENT APPLICATION NUMBER: US/09/814,604
CURRENT FILING DATE: 2001-03-22
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-09-814-604-26
Query Match 1.4%; Score 7; DB 11; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 FTGLSLIA 82
Db 76 FTGLSLIA 82

Db 8 FTGLSIA 14

```

RESULT 15
US-09-922-226-12
; Sequence 12, Application US/09922226
; Publication No. US20030077664A1
; GENERAL INFORMATION:
; APPLICANT: Zhao, Yi
; APPLICANT: Thacher, Scott M.
; APPLICANT: Xiao, Jia-Hao
; APPLICANT: Kusari, Jyotirmoy
; APPLICANT: Chandraratna, Roshanra A.
; TITLE OF INVENTION: Methods of Screening For Compounds That
; FILE REFERENCE: P-AR 4681
; CURRENT APPLICATION NUMBER: US/09/922,226
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/284,797
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-226-12
    
```

```

Query Match 1.4%; Score 7; DB 11; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 FTGLSIA 82
DB 8 FTGLSIA 14
    
```

Search completed: November 21, 2003, 14:29:34
 Job time : 37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2003, 02:08:36 ; Search time 8288 Seconds

(without alignments) 11219.544 Million cell updates/sec

Title: US-09-938-803-26

Perfect score: 2273

Sequence: 1 ggggggtgaagcagcagctt.....ttttcaaaaaaaaaaaaaa 2273

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ses:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ses:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hugo_hum:*
40: em_hugo_mus:*
41: em_hugo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2273	100.0	2273	6	AX048067 Sequence
2	2241.2	98.6	2264	6	AX073570 Sequence
3	2215	97.4	2482	6	AK096775 Homo sapi
4	2045.4	90.0	2058	6	BC000664 Homo sapi
5	1842	81.0	1842	6	AX331926 Sequence
6	1842	81.0	1842	6	AX410769 Sequence
7	1842	81.0	1842	6	HSU72515 Sequence
8	1461	64.3	1461	6	AX073560 Sequence
9	1388.2	61.1	1573	6	AK058063 Homo sapi
10	1240.4	54.6	2128	10	AK098090 Mus muscu
11	1233.4	54.3	1898	10	BC006753 Mus muscu
12	1229.4	54.1	1968	10	AY028317 Mus muscu
13	1143.4	50.3	1146	9	BT007000 Homo sapi
14	1142.4	50.3	1146	12	BT007735 Synthetic
15	708.6	31.2	155975	9	AC006512 Homo sapi
16	708.6	31.2	222930	9	HSU47924 Human chrom
17	433.2	19.1	450	6	BD023986 Sequence
18	425.4	18.7	1641	6	BC009856 Homo sapi
19	254.8	11.2	990	6	AX331531 Sequence
20	254.8	11.2	990	6	AX331927 Sequence
21	254.8	11.2	990	6	AX335065 Sequence
22	249.6	11.0	227538	10	AC002397 Human alter
23	249.6	11.0	227538	10	AC002397 Mouse chr
24	236.4	10.4	169660	9	AY050668 Homo sapi
25	236.4	10.4	169661	9	AC011479 Homo sapi
26	235.6	10.4	192440	9	AL139316 Human chr
27	233.2	10.3	176541	9	AL929325 Human DNA
28	233	10.3	140348	9	AL390725 Human DNA
29	233	10.3	321519	2	AL1714004 Human sapi
30	232.8	10.2	147054	2	AL356582 Homo sapi
31	232.4	10.2	67885	2	AC107924 Homo sapi
32	232.4	10.2	86156	2	AL162262 Homo sapi
33	232.4	10.2	115093	9	AC091558 Homo sapi
34	232.2	10.2	146350	9	HS225115 Homo sapi
35	232.2	10.2	340000	9	HS21C100 Homo sapi
36	232	10.2	321519	2	AL1714004 Homo sapi
37	231.2	10.2	167548	9	AL391809 Human DNA
38	231	10.2	128829	9	AC127002 Homo sapi
39	231	10.2	147419	9	AL031736 Human DNA
40	231	10.2	194296	9	AL354864 Human DNA
41	231	10.2	201508	2	AC026290 Homo sapi
42	230.8	10.2	37680	2	AL359985 Homo sapi
43	230.8	10.2	93582	2	AC090681 Homo sapi
44	230.8	10.2	146017	2	AC027473 Homo sapi
45	230.8	10.2	151465	2	AC117371 Homo sapi

ALIGNMENTS

RESULT 1
AX048067
LOCUS AX048067 2273 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 61 from Patent WO0070047.
ACCESSION AX048067
VERSION AX048067.1 GI:11876890
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Yue,H., Tang,Y.T., Lai,P., Reddy,R., Batra,S., Baughn,M.R.,
Yang,J., Azimzai,Y., Lu,D.A., Au-Young,J. and Shih,L.L.
Full-length molecules expressed in human tissues

JOURNAL Patent: WO 0070047-A 61.23-NOV-2000;
Incyle Genomics, Inc. (US)
Location/Qualifiers
FEATURES
source 1..2273
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyle ID No: 1928920CB1"
BASE COUNT 494 a 630 c 547 g 602 t
ORIGIN
Query Match 100.0%; Score 2273; DB 6; Length 2273;
Best Local Similarity 100.0%; Pred. No.0;
Matches 2273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGGTGAAGCGATACGTTTGGCCCGCATTCGGGGCGCGCGACTGGGGGGTCCCTGTG 60
DB 1 GGGGGTGAAGCGATACGTTTGGCCCGCATTCGGGGCGCGCGACTGGGGGGTCCCTGTG 60
QY 61 GGGCTCCCGAGTTAAGATGGCGTCTCAGCGAGGGGGAGAGAGGGACTGTGGTGGCG 120
DB 61 GGGCTCCCGAGTTAAGATGGCGTCTCAGCGAGGGGGAGAGAGGGACTGTGGTGGCG 120
QY 121 TGGCGGGGGTTTGCAGTGGGGTTTCCAGGAGCGTGAACCTTAACAAGTTGGCGACGTC 180
DB 121 TGGCGGGGGTTTGCAGTGGGGTTTCCAGGAGCGTGAACCTTAACAAGTTGGCGACGTC 180
QY 181 TGGCGCGGTGAGAAAGAGCGCGTGGCGTGCATCTTCATCTTCTGGGTTACCCCTTGG 240
DB 181 TGGCGCGGTGAGAAAGAGCGCGTGGCGTGCATCTTCATCTTCTGGGTTACCCCTTGG 240
QY 241 CTTTGTGTTTATGCGCATTAACCTTTTCTACAAGAGACTTACCTTACCTCTTCCATA 300
DB 241 CTTTGTGTTTATGCGCATTAACCTTTTCTACAAGAGACTTACCTTACCTCTTCCATA 300
QY 301 CCTTACAGGCGCTCTCAATGCTTATTTTAACCTTTGGAACCAAGCTCTACACCTCCGTC 360
DB 301 CCTTACAGGCGCTCTCAATGCTTATTTTAACCTTTGGAACCAAGCTCTACACCTCCGTC 360
QY 361 TGTGTATGTGCTTCAGATTCCTCATCTTGAAGTGGCGCGACCATCATCGCGTCC 420
DB 361 TGTGTATGTGCTTCAGATTCCTCATCTTGAAGTGGCGCGACCATCATCGCGTCC 420
QY 421 TCACTACCTTTTGGCTTCCAGATGGCTTCTGGTGGATCTATTAACCTGCCACCG 480
DB 421 TCACTACCTTTTGGCTTCCAGATGGCTTCTGGTGGATCTATTAACCTGCCACCG 480
QY 481 GCAACTACGATATCAAGTGAAGTGCACATTTGTCTGACTTTGAAGCTGATTGGT 540
DB 481 GCAACTACGATATCAAGTGAAGTGCACATTTGTCTGACTTTGAAGCTGATTGGT 540
QY 541 TGGCTGTGACTTCTTGAAGAGAGAAAGATCAGAAATTCCTGTCTGAGCAACAGA 600
DB 541 TGGCTGTGACTTCTTGAAGAGAGAAAGATCAGAAATTCCTGTCTGAGCAACAGA 600
QY 601 AATATGCCAATGCTGGTCTTCTCCCTGGTGAAGTGGTGGTCTCTACTTCTATG 660
DB 601 AATATGCCAATGCTGGTCTTCTCCCTGGTGAAGTGGTGGTCTCTACTTCTATG 660
QY 661 GGGCTCTTGTGAGGGCCCAAGTCTCAATGATCACTACATGAAGCTGTGAGGAG 720
DB 661 GGGCTCTTGTGAGGGCCCAAGTCTCAATGATCACTACATGAAGCTGTGAGGAG 720
QY 721 AGCTGATTGACATACAGAAAGATACAAACAGCATCATTCCTGCTCAAGCGCTGA 780
DB 721 AGCTGATTGACATACAGAAAGATACAAACAGCATCATTCCTGCTCAAGCGCTGA 780
QY 781 GTCTGGGCTTTTCTACTAGTGGGTACACATGCTCAGCCCGACATACAGAAAGCT 840
DB 781 GTCTGGGCTTTTCTACTAGTGGGTACACATGCTCAGCCCGACATACAGAAAGCT 840
QY 841 ATCTCCTCACTGAAGATATAGCAACACCCCTTCTGGTTCGGGTGATATAGTGA 900
DB 841 ATCTCCTCACTGAAGATATAGCAACACCCCTTCTGGTTCGGGTGATATAGTGA 900

DB 841 ATCTCCTCACTGAAGATATAGCAACACCCCTTCTGGTTCGGGTGATATAGTGA 900
QY 901 TCTGGGGCAAGTTTGTGCTGTACAATATATGCACTGTTGGTGTACAGAGAGATAT 960
DB 901 TCTGGGGCAAGTTTGTGCTGTACAATATATGCACTGTTGGTGTACAGAGAGATAT 960
QY 961 GCATTTTGAAGGGCCCTGCTTCAATGCTTTGAAGAAAAAGGCAAGCAAGTGGGATG 1020
DB 961 GCATTTTGAAGGGCCCTGCTTCAATGCTTTGAAGAAAAAGGCAAGCAAGTGGGATG 1020
QY 1021 CCTGTGCAACATGAAGTGTGGCTCTTGAAGAAACCAACCCCGCTTCACTGGACCATTTG 1080
DB 1021 CCTGTGCAACATGAAGTGTGGCTCTTGAAGAAACCAACCCCGCTTCACTGGACCATTTG 1080
QY 1081 CCTCATTTCAACATCAACACCAACCGCTGGGTGGCCGCTACATCTTCAACGACTCAAGT 1140
DB 1081 CCTCATTTCAACATCAACACCAACCGCTGGGTGGCCGCTACATCTTCAACGACTCAAGT 1140
QY 1141 TCTTGGAAATTAAGAACTCTCAGAGTCTCTGTTGCTATTTCTGGGCTCTGGCAG 1200
DB 1141 TCTTGGAAATTAAGAACTCTCAGAGTCTCTGTTGCTATTTCTGGGCTCTGGCAG 1200
QY 1201 GCTGCACTCAGGATACCTGTGCTCTGCTTCCAGATGGAATTCCTCATTTGTTATGGAAA 1260
DB 1201 GCTGCACTCAGGATACCTGTGCTCTGCTTCCAGATGGAATTCCTCATTTGTTATGGAAA 1260
QY 1261 GACAGGCTCCAGGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCCGCTATTA 1320
DB 1261 GACAGGCTCCAGGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCCGCTATTA 1320
QY 1321 TCTTCCAGGCTTCTATTTTGGTGAACAACAACCACTGAGGCTCTTCAAGGTTACT 1380
DB 1321 TCTTCCAGGCTTCTATTTTGGTGAACAACAACCACTGAGGCTCTTCAAGGTTACT 1380
QY 1381 CCATGACTGCTTCTGCTCTTCACTGAGGAGCAATAGCTTAAGTGTATTAATCCATCT 1440
DB 1381 CCATGACTGCTTCTGCTCTTCACTGAGGAGCAATAGCTTAAGTGTATTAATCCATCT 1440
QY 1441 ATTTCTTGGCAATCTTCTCTGAGCTTCAATTCATTTGCTTATTAATCACAAG 1500
DB 1441 ATTTCTTGGCAATCTTCTCTGAGCTTCAATTCATTTGCTTATTAATCACAAG 1500
QY 1501 CAATGTGCAAGAAAGAAAGTAAAGAAATGAATGAATTCATTTCCCTGGTGGCT 1560
DB 1501 CAATGTGCAAGAAAGAAAGTAAAGAAATGAATGAATTCATTTCCCTGGTGGCT 1560
QY 1561 GTGCGGAACTGGTGCAGAAACTACTGCTCCCTTTCAACAGCACTCTTGGCCGAG 1620
DB 1561 GTGCGGAACTGGTGCAGAAACTACTGCTCCCTTTCAACAGCACTCTTGGCCGAG 1620
QY 1621 CAGAGATGAAAAGCAGAGAGTGAAGATGAGTTCAGAGCTGTGCTCTGCTGCG 1680
DB 1621 CAGAGATGAAAAGCAGAGAGTGAAGATGAGTTCAGAGCTGTGCTCTGCTGCG 1680
QY 1681 AGCAAGTCTTCAATTTGGGGCCAAAAGGGGAACTTTTGGAGAAAGGCGTTCGCTT 1740
DB 1681 AGCAAGTCTTCAATTTGGGGCCAAAAGGGGAACTTTTGGAGAAAGGCGTTCGCTT 1740
QY 1741 GTTACCCACGCTGGAATGAGTGGCGGATCTCAGCTCAACCGAACCTCCTCTGGG 1800
DB 1741 GTTACCCACGCTGGAATGAGTGGCGGATCTCAGCTCAACCGAACCTCCTCTGGG 1800
QY 1801 TTCAAGTATTTTCTGCTCAGCTCCCAAGTATGTTGGAAATACAGGCAACGACCATG 1860
DB 1801 TTCAAGTATTTTCTGCTCAGCTCCCAAGTATGTTGGAAATACAGGCAACGACCATG 1860
QY 1861 CCAGCTAATTTTGTATTTTCAAGTGAAGAGGAAATTCACACGTTGGCGAGGCTGTC 1920
DB 1861 CCAGCTAATTTTGTATTTTCAAGTGAAGAGGAAATTCACACGTTGGCGAGGCTGTC 1920
QY 1921 TCGAATCTCTGACCGGAAGTATCAACCGGCTCTCGGCTCCCAAGTGTGGGATTAACAG 1980
DB 1921 TCGAATCTCTGACCGGAAGTATCAACCGGCTCTCGGCTCCCAAGTGTGGGATTAACAG 1980

QY 1981 GCGTAGCCACCGTCCCGGCCCCAAAGGGGAAACTCTGTGGAGAGAGAGAGGGGCTCA 2040
Db 1981 GCGTAGCCACCGTCCCGGCCCCAAAGGGGAAACTCTGTGGAGAGAGAGAGGGGCTCA 2040
QY 2041 CATCTCCCTCTGATTCCTCCCATGACATTCGCTTATCTCTCCCATCTAGCCAGAAATC 2100
Db 2041 CATCTCCCTCTGATTCCTCCCATGACATTCGCTTATCTCTCCCATCTAGCCAGAAATC 2100
QY 2101 TATCTGTTTTTCTCTGCAATTTAATGATTTGATTTGCGGCTTACCAACACCCCC 2160
Db 2101 TATCTGTTTTTCTCTGCAATTTAATGATTTGATTTGCGGCTTACCAACACCCCC 2160
QY 2161 CCCATGGGGGGGTGAGAGAGGGGTGCAAGGCCCTGCTGCTCACTTTTCTACCTTGGAA 2220
Db 2161 CCCATGGGGGGGTGAGAGAGGGGTGCAAGGCCCTGCTGCTCACTTTTCTACCTTGGAA 2220
QY 2221 CTGATTTAGATTAATCACTTCTGTTTCTTCAAGTTTTCACAAAAA 2273
Db 2221 CTGATTTAGATTAATCACTTCTGTTTCTTCAAGTTTTCACAAAAA 2273

RESULT 2
AX073570 2264 bp DNA linear PAT 06-FEB-2001
LOCUS AX073570
DEFINITION Sequence 22 from Patent WO0104297.
ACCESSION AX073570
VERSION AX073570.1 GI:12709984
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Kato, S. and Kimura, T.
Human proteins having hydrophobic domains and dnas encoding these proteins
Patent: WO 0104297-A 22 18-JAN-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)

FEATURES
source location/Qualifiers

CDS
1..2264
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
85..1548
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28402.1"
/translation="MASSAEDEGTVAAGVLAGVLOSFOELSLNKLATSLGASEOARL
IISIFLGPALFVRHLYPKETVLIHLEHPTGSLIAYFNGNDLHSLICTIVQPL
ILRLMGRITAVLTTCFQOMAYLAGVYTTATGNDIMTHCHVLIHLGLADYF
DGGKONSLSBQKRAIRGVPSLLEVAQSFYFAGLVGPQFSNMHWKLVQGLID
IPGKIPNSIIIPALRSLGLFVLGVYTLSPHITEDYLTEDYDNHPPFMVYLIM
GKFLVYKVTICVLVEGVCIILGLFNGFEKAKMDKAMKVMLEFETNRFGTI
ASFNINNTAMARVIFKRLKFLGNKELSGSLFLALMHGHSYLVCFOPEFLIVI
VEROARLIQSPITSLKLAITVLOPBYLIQOQTHMLFMGSMATACFLFTMDKLVK
YKSTFLCHIFLSLFLPIIPLHFKAMVRKELKAKME"

BASE COUNT 482 a 629 c 550 g 603 t
ORIGIN

Query Match 98.6%; Score 2241.2; DB 6; Length 2264;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2254; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 GGGGGTGAAGCATA-CGTTTTCCTCGCATTCGGGGCGCGGAGCTGGGGGGTCCCTGT 59
Db 7 GGGGGTGAAGCATAAGCGTTTTCCTCGCATTCGGGGCGCGGAGCTGGGGGGTCCCTGT 66
QY 60 GGGGGTCCCGAGTAAAGATGGCGTCTCAGCGGAGGGAGAGAGAGAGAGAGAGAGAG 119
Db 67 GGGGGTCCCGAGTAAAGATGGCGTCTCAGCGGAGGGAGAGAGAGAGAGAGAGAGAG 126

QY 120 CTGGCGGGGGTCTTCAGTCCGGGTTTCCAGAGAGTGAACCTTAAACAAGTTGGCAGCTCC 179
Db 127 CTGGCGGGGGTCTTCAGTCCGGGTTTCCAGAGAGTGAACCTTAAACAAGTTGGCAGCTCC 186
QY 180 CTGGCGGGGTCAAGAACAGGGCGCTGGGCTGATTCATCTCATCTTCTGGGTTTACCCTTT 239
Db 187 CTGGCGGGGTCAAGAACAGGGCGCTGGGCTGATTCATCTCATCTTCTGGGTTTACCCTTT 246
QY 240 GCTTTGTTTATTCGCAATACCTTTTCAAGAGAGACCTTACCTTACCTTCCAT 299
Db 247 GCTTTGTTTATTCGCAATACCTTTTCAAGAGAGACCTTACCTTACCTTCCAT 306
QY 300 ACCTTTACAGGCTCTCAATGCTTATTTTAACTTTTGAACACAGCTTACACCTCCG 359
Db 307 ACCTTTACAGGCTCTCAATGCTTATTTTAACTTTTGAACACAGCTTACACCTCCG 366
QY 360 CTGTGATTTGCTTCACTTCCTCAATCCTTGCATTAATGAGGCGGACCATCATCTGCTC 419
Db 367 CTGTGATTTGCTTCACTTCCTCAATCCTTGCATTAATGAGGCGGACCATCATCTGCTC 426
QY 420 CTCACTACCTTTTGTTCAGATAGGCTTACCTTCTGCTGATTAATTAATTAACCTGCCAC 479
Db 427 CTCACTACCTTTTGTTCAGATAGGCTTACCTTCTGCTGATTAATTAATTAACCTGCCAC 486
QY 480 GCAACTACGATATCAAGTGAACAAATGACACATTTGTCTGACTTGAAGCTGATTTGAT 539
Db 487 GCAACTACGATATCAAGTGAACAAATGACACATTTGTCTGACTTGAAGCTGATTTGAT 546
QY 540 TTGGCTGTTGACTATTGACGAGAGGAAAGATCAAGATTCCTTGTCTTGAAGCAG 599
Db 547 TTGGCTGTTGACTATTGACGAGAGGAAAGATCAAGATTCCTTGTCTTGAAGCAG 606
QY 600 AAATATGCAATACGATGATTTCTTCCCTGCTGAGAGTTGCTGATTTCTCTCACTTAT 659
Db 607 AAATATGCAATACGATGATTTCTTCCCTGCTGAGAGTTGCTGATTTCTCTCACTTAT 666
QY 660 GGGGCTCTTGTGATGAGGCGCCAGTTCATTAATGAATCAATGATGATGATGATGATGAT 719
Db 667 GGGGCTCTTGTGATGAGGCGCCAGTTCATTAATGAATCAATGATGATGATGATGATGAT 726
QY 720 GAGCTGATTAATACATACAGAAAGATTAACCAACAGATCAATTCCTGCTCAAGGCTGT 779
Db 727 GAGCTGATTAATACATACAGAAAGATTAACCAACAGATCAATTCCTGCTCAAGGCTGT 786
QY 780 AGCTGGGCTTTTCTTACTTGTGGGCTTACACACTGCTCAGGCCCAATACACAAAGAC 839
Db 787 AGCTGGGCTTTTCTTACTTGTGGGCTTACACACTGCTCAGGCCCAATACACAAAGAC 846
QY 840 TATCTCCTCACTGAAGATATGACCAACACCCCTTCTGATTCGCTGATGATGATGATGAT 899
Db 847 TATCTCCTCACTGAAGATATGACCAACACCCCTTCTGATTCGCTGATGATGATGATGAT 906
QY 900 ATCTGGGCAAGTTTGTCTGTACAAATATGATCACTTGTGCTGTGATGATGATGATGAT 959
Db 907 ATCTGGGCAAGTTTGTCTGTACAAATATGATCACTTGTGCTGTGATGATGATGATGAT 966
QY 960 TGCATTTGAAGGGGCTGGGCTTCAATGGCTTTGAAGAAAGGCAAGCAAGTGGAT 1019
Db 967 TGCATTTGAAGGGGCTGGGCTTCAATGGCTTTGAAGAAAGGCAAGCAAGTGGAT 1026
QY 1020 GCTGTGCAACATGAAGTGTGGCTTTTGAACCAACCCCGCTTCACTGGACCAAT 1079
Db 1027 GCTGTGCAACATGAAGTGTGGCTTTTGAACCAACCCCGCTTCACTGGACCAAT 1086
QY 1080 GCTTCAATCAACATCAACCAACGCTGGGTGGCGCTACATCTTCAAAAGCACTCAAG 1139
Db 1087 GCTTCAATCAACATCAACCAACGCTGGGTGGCGCTACATCTTCAAAAGCACTCAAG 1146
QY 1140 TTTCTTGAATTAAGAACTCTTCAAGGCTCTCTGTTGATTTCTGAGGCTTGGGAC 1199
Db 1147 TTTCTTGAATTAAGAACTCTTCAAGGCTCTCTGTTGATTTCTGAGGCTTGGGAC 1206
QY 1200 GGCCTGCACTAGATACCTGCTTCTTCAAGATGATTAATCTCATGTTATTTGTGAA 1259

Db 1207 GGGCTGACCTCAGAGTAACCTGGTCTCTTCAGATGGAATCTCTCATGTATTGGGAA 1266
 Qy 1260 AGACAGAGCTGCCAGGCTCATTTCAAGAGAGAGCCCACTGAGCAAGCTGGCCGCTTACT 1319
 Db 1267 AGACAGAGCTGCCAGGCTCATTTCAAGAGAGAGCCCACTGAGCAAGCTGGCCGCTTACT 1326
 Qy 1320 GTTCTCAGAGCCCTTCTACTATTGTGTGCAAGACATCATCTGCTCTTCAAGGTTAC 1379
 Db 1327 GTTCTCAGAGCCCTTCTACTATTGTGTGCAAGACATCATCTGCTCTTCAAGGTTAC 1386
 Qy 1380 TCCATACAGCTGCTTCTGCTCTTCAAGTGGGCAAAATGCTTAAGGTATTAATCCATC 1439
 Db 1387 TCCATACAGCTGCTTCTGCTCTTCAAGTGGGCAAAATGCTTAAGGTATTAATCCATC 1446
 Qy 1440 TATTTCTTGGCCACATCTTCTCTGAGCTCATTTCAATATTTGCTTATTTCAAAA 1499
 Db 1447 TATTTCTTGGCCACATCTTCTCTGAGCTCATTTCAATATTTGCTTATTTCAAAA 1506
 Qy 1500 GCATAGGTGCAAGAAAGAAAGTAAAGAGATGGAATATTCATTTCTCTGTGGCC 1559
 Db 1507 GCATAGGTGCAAGAAAGAAAGTAAAGAGATGGAATATTCATTTCTCTGTGGCC 1566
 Qy 1560 TGTGCGGAGCTGGTGAAGAAACTACTGCTCTCTTTTCAAGACACTCTTGGCCCAAG 1619
 Db 1567 TGTGCGGAGCTGGTGAAGAAACTACTGCTCTCTTTTCAAGACACTCTTGGCCCAAG 1626
 Qy 1620 GCAGAGATGGAAGAAAGCAAGGAGTGAAGATGATGCTTCCAGCTGTGCTCTGCTG 1679
 Db 1627 GCAGAGATGGAAGAAAGCAAGGAGTGAAGATGATGCTTCCAGCTGTGCTCTGCTG 1686
 Qy 1680 CAGCCAACTTCTATTTGGGGCCAAAGGGAACTTTTGTGAGAAAGGCTTGTCTT 1739
 Db 1687 CAGCCAACTTCTATTTGGGGCCAAAGGGAACTTTTGTGAGAAAGGCTTGTCTT 1746
 Qy 1740 TGTACCAAGCTGGAATGAGTGGGGGATCTCAGCTCACCGCAACTCTCCTGCTG 1799
 Db 1747 TGTACCAAGCTGGAATGAGTGGGGGATCTCAGCTCACCGCAACTCTCCTGCTG 1806
 Qy 1800 GTTCAAGTATTTTCTGCTCAGCTCTCCCAAGTACTGGAATTAACAGGAGCCCAACAT 1859
 Db 1807 GTTCAAGTATTTTCTGCTCAGCTCTCCCAAGTACTGGAATTAACAGGAGCCCAACAT 1866
 Qy 1860 GCGCAGCTATTTTGTATTTTCAATGAAACGGGATTTTCAACAGTGTGGCCAGGCTGT 1919
 Db 1867 GCGCAGCTATTTTGTATTTTCAATGAAACGGGATTTTCAACAGTGTGGCCAGGCTGT 1926
 Qy 1920 CTGCAACTCTGACCGCAAGTATTCACCGGCTCCGCTCCCAAGTGTGGGATTTACA 1979
 Db 1927 CTGCAACTCTGACCGCAAGTATTCACCGGCTCCGCTCCCAAGTGTGGGATTTACA 1986
 Qy 1980 GGGGTGAGCCACGCTGCGCCGCCAAAGGGGAACTTTTGTGAGAGAGAGAGGGCTC 2039
 Db 1987 GGGGTGAGCCACGCTGCGCCGCCAAAGGGGAACTTTTGTGAGAGAGAGAGGGCTC 2046
 Qy 2040 ACATCTTCTCTGATTTTCCCAATGCAATGCTTATCTCTCCCATCTAGCCAGGAAT 2099
 Db 2047 ACATCTTCTCTGATTTTCCCAATGCAATGCTTATCTCTCCCATCTAGCCAGGAAT 2106
 Qy 2100 CTATGTGTTTCTTCTGCAATTTTACTATGATGTGTGTCGCTACCAACACCC 2159
 Db 2107 CTATGTGTTTCTTCTGCAATTTTACTATGATGTGTGTCGCTACCAACACCC 2166
 Qy 2160 CCCCATGAGGGGGTGAAGAGGGGTGAAGGCCCTGCTGCTCACTTTTCTACTTTGA 2219
 Db 2167 CCCCATGAGGGGGTGAAGAGGGGTGAAGGCCCTGCTGCTCACTTTTCTACTTTGA 2226
 Qy 2220 ACTGTATTAATAAATCACTTCTGTTTGTTCAGTTT 2257
 Db 2227 ACTGTATTAATAAATCACTTCTGTTTGTTCAGTTT 2264

RESULT 3

AK096775 2482 bp mRNA linear pri 15-JUL-2002
 LOCUS AK096775 Homo sapiens cDNA FLJ39456 fis, clone PROST2010782, highly similar
 DEFINITION to Human C3f mRNA.
 ACCESSION AK096775
 VERSION AK096775.1 GI:21756342
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS
 1 Nimomiya, K., Magatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
 Kodaira, H., Furuya, T., Takahashi, M., Kikawa, E., Omura, Y., Abe, K.,
 Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, J., Isono, Y., Kawai-Hiro, Y., Saito, K., Nishikawa, T.,
 Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
 Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2482)
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genominfo@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan: cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5- & 3- end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source
 1..2482
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PROST2010782"
 /issue_type="Prostate"
 /clone_id="PROST2"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 526 a 687 c 615 g 654 t
 ORIGIN
 Query Match 97.4%; Score 2215; DB 9; Length 2482;
 Best Local Similarity 99.6%; Pred. No. 0; Indels 3; Gaps 3;
 Matches 2251; Conservative 0; Mismatches 5;
 Qy 1 GGGGGTGAAGAGCATTA-CGTTTGGCCGCAATTCGGGGGCGCGGAGTGGGGGGTCCCTGT 59
 Db 226 GGGGGTGAAGAGCATTAAGCTTTTCCCGCATTCGGGGGCGCGGAGTGGGGGGTCCCTGT 285
 Qy 60 GGGGCTCCGAGTAAAGATGCGTCTCAGCGGAGGGGAGCAGAGGGGACTGTGTGGCG 119
 Db 286 GGGGCTCCGAGTAAAGATGCGTCTCAGCGGAGGGGAGCAGAGGGGACTGTGTGGCG 345
 Qy 120 CTGGCGGGGGTCTGAGATCGGGTTTCCAGAGCTGAGCTTAAACAAGTTGGGAGATGC 179
 Db 346 CTGGCGGGGGTCTGAGATCGGGTTTCCAGAGCTGAGCTTAAACAAGTTGGGAGATGC 405
 Qy 180 CTGGGCGCTCAAGACAGCGCGTGCAGCTGATCATCTCATCTTCTGGGTTACCCCTTT 239
 Db 406 CTGGGCGCTCAAGACAGCGCGTGCAGCTGATCATCTCATCTTCTGGGTTACCCCTTT 465
 Qy 240 GCTTTGTTTATTCGACATTAACCTTTTCTAACAGAGACCTTACCTCACTTCACCTTTGCAT 299
 Db 466 GCTTTGTTTATTCGACATTAACCTTTTCTAACAGAGACCTTACCTCACTTCACCTTTGCAT 525

QY 300 ACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGGAAACAGCTTACCACTCCCTG 359
DB 526 ACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGGAAACAGCTTACCACTCCCTG 585
QY 360 CTGTGATTTGCTTCAAGTTCTCTCATCTTCCATGATGAGGCGGCACTCACTGCGCTG 419
DB 586 CTGTGATTTGCTTCAAGTTCTCTCATCTTCCATGATGAGGCGGCACTCACTGCGCTG 645
QY 420 CTCACTACCTTTGCTTCCAGATGAGCTTCTTGGCTGGATGATGATGATGATGATGATG 479
DB 646 CTCACTACCTTTGCTTCCAGATGAGCTTCTTGGCTGGATGATGATGATGATGATGATG 705
QY 480 GGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 539
DB 706 GGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 765
QY 540 TTGGCTGTTGATCTATTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 599
DB 766 TTGGCTGTTGATCTATTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 825
QY 600 AATATAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
DB 826 AATATAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 885
QY 660 GGGGCTTCTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 719
DB 886 GGGGCTTCTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 945
QY 720 GAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779
DB 946 GAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005
QY 780 AGCTGGGCTTTTCTACCTAGGAGCTACACAGCTGCTGAGGCTGAGGCTGAGGCTGAGG 839
DB 1006 AGCTGGGCTTTTCTACCTAGGAGCTACACAGCTGCTGAGGCTGAGGCTGAGGCTGAGG 1065
QY 840 TATCTCTCTACCTAGGAGCTACACAGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 899
DB 1066 TATCTCTCTACCTAGGAGCTACACAGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1125
QY 900 ATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 959
DB 1126 ATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1185
QY 960 TGCATTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1019
DB 1186 TGCATTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1245
QY 1020 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1079
DB 1246 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1305
QY 1080 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1139
DB 1306 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1365
QY 1140 TTTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1199
DB 1366 TTTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1424
QY 1200 GGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
DB 1425 GGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1484
QY 1260 AAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1319
DB 1485 AAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1544
QY 1320 GTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1379
DB 1545 GTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1604
QY 1380 TCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1439

DB 1605 TCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1664
QY 1440 TATTTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1499
DB 1665 TATTTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1724
QY 1500 GCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1559
DB 1725 GCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1784
QY 1560 TGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1619
DB 1785 TGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1844
QY 1620 GCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1679
DB 1845 GCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1904
QY 1680 CAGCCAGGCTTCAATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1739
DB 1905 CAGCCAGGCTTCAATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1963
QY 1740 TGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1799
DB 1964 TGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2023
QY 1800 GTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1859
DB 2024 GTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2083
QY 1860 GCCCAGCTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1919
DB 2084 GCCCAGCTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2143
QY 1920 CTGGAAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1979
DB 2144 CTGGAAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2203
QY 1980 GGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2039
DB 2204 GGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2263
QY 2040 ACATCTCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
DB 2264 ACATCTCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2323
QY 2100 CTATTTGTTTTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2159
DB 2324 CTATTTGTTTTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2383
QY 2160 CCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2219
DB 2384 CCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2443
QY 2220 ACTGATTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2258
DB 2444 ACTGATTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2482

RESULT 4
LOCUS BC000664 2058 bp mRNA linear PRI 22-APR-2003
DEFINITION Homo sapiens putative protein similar to neesby (Drosophila), mRNA
(cDNA clone MGC:1311 IMAGE:349388), complete cds.
ACCESSION BC000664
VERSION BC000664.1 GI:12653756
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2058)

AUTHORS	
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Stemann, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Hsieh, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Bhat, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Szechenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
TITLE	
human and mouse cDNA sequences	
JOURNAL	
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
MEDLINE	
22388257	
PUBMED	
12477932	
JOURNAL	
2 (bases 1 to 2058)	
AUTHORS	
Straussberg, R.	
TITLE	
Submitted (15-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	
NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT	
Contact: MGC help desk Email: cgapsb@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIND) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Ahner, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hashiguchi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaesi, R., Maduro, Q.L., Mastello, C., Maekari, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stancik, S., Thomas, P.J., Touchman, J.W., Tsurgan, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggin, L., Young, A., Zhang, L.-H. and Green, E.D.	
FEATURES	
Source	
1. .2058 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:1311 IMAGE:3349388" /tissue_type="Colon, adenocarcinoma" /clone_id="NIH_MGC_15" /lab_host="DH10B-R" /note="Vector: pOTB7" 1. .2058 /gene="C3F" /db_xref="LocusID:10162" 177. .1322 /codon_start=1 /product="putative protein similar to nesy" /protein_id="AAH00664.1" /db_xref="GI:12653757"	
FEATURES	
Location/Qualifiers	
1. .2058 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:1311 IMAGE:3349388" /tissue_type="Colon, adenocarcinoma" /clone_id="NIH_MGC_15" /lab_host="DH10B-R" /note="Vector: pOTB7" 1. .2058 /gene="C3F" /db_xref="LocusID:10162" 177. .1322 /codon_start=1 /product="putative protein similar to nesy" /protein_id="AAH00664.1" /db_xref="GI:12653757"	
FEATURES	
Location/Qualifiers	
1. .2058 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:1311 IMAGE:3349388" /tissue_type="Colon, adenocarcinoma" /clone_id="NIH_MGC_15" /lab_host="DH10B-R" /note="Vector: pOTB7" 1. .2058 /gene="C3F" /db_xref="LocusID:10162" 177. .1322 /codon_start=1 /product="putative protein similar to nesy" /protein_id="AAH00664.1" /db_xref="GI:12653757"	

/db_xref="LocusID:10162"
 /translation="MGRITPAVLTTPFCROMAYLAGYYTATGANDIKMTPHCVLT
 KLIGLAVIDFDGKQKONSISSEQQRALRGVSLLELVAGFYPGAFVLPQPSMNH
 MLYOGELVIDPFGKIPNSIIPALKSLSPFLVYTLSPITADYLTEDYDNHPE
 WRCWMLIMGKFLVYKYVTCVLVEGCIILGLFNGEESGKAKMDACAMKVLV
 EIRPPTGIAEFNITNMAWRYI PKRLKFIKNSLSGLFLFALHGHSGVLV
 CROMFLIVIVROARLQESPTLSKLAIVIVLPFYVLTNOOTIHWLMGMSMAFC
 LFTWKMVKVYSITFLGHIFLSLFLPIPIHKMVRKELKKNNE"
 BASE COUNT 470 a 575 c 457 g 556 t
 ORIGIN
 Query Match 90.0%; Score 2045.4; DB 9; Length 2058;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2046; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

227 GGGTTACCCCTTGGCTTTGTTTATCGGATTACCTTTCTACAGAGACCTACTCAT 286
 8 GGGTTACCCCTTGGCTTTGTTTATCGGATTACCTTTCTACAGAGACCTACTCAT 67
 287 CCACCTCTTCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGGAACCAAGCT 346
 68 CCACCTCTTCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGGAACCAAGCT 127
 347 CTACCACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406
 128 CTACCACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 187
 407 CATCACTGCGCTCTCTCACTACCTTTTGTCTTCAGATGAGCTTACTTGGCTGATAGTA 466
 188 CATCACTGCGCTCTCTCACTACCTTTTGTCTTCAGATGAGCTTACTTGGCTGATAGTA 247
 467 TTACATCTGCAACCGGCACTAGATATCAAGTGAACAAGCCATTTGTTTGTACTTT 526
 248 TTACATCTGCAACCGGCACTAGATATCAAGTGAACAAGCCATTTGTTTGTACTTT 307
 527 GAAGCTGATGTTGGTGGCTGTTGACTACTTGTGACGAGGAGAAAGATCAAAATCTTGTG 586
 308 GAAGCTGATGTTGGTGGCTGTTGACTACTTGTGACGAGGAGAAAGATCAAAATCTTGTG 367
 587 CTCTGAGCAACAGAAATATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
 368 CTCTGAGCAACAGAAATATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
 647 CTCTGATCTTATGAGGCTCTTGTGTAGGCGCCGCTTCAATGAAATCACTACATGAA 706
 428 CTCTGATCTTATGAGGCTCTTGTGTAGGCGCCGCTTCAATGAAATCACTACATGAA 487
 707 GCTGTGACAGGAGAGCTGATGACATTAACAGGAAAGATACCAACAGATCTTCTGCG 766
 488 GCTGTGACAGGAGAGCTGATGACATTAACAGGAAAGATACCAACAGATCTTCTGCG 547
 767 TCTCAAGCGCTGAGTCTGCGCTTTTCTAAGTGGCTACACTGCTACGCCCA 826
 548 TCTCAAGCGCTGAGTCTGCGCTTTTCTAAGTGGCTACACTGCTACGCCCA 607
 827 CATCAAGAGAGATATCTCTCACTAAGCTATGACCAACCCCTTGTGCTGCTGCTGCTG 886
 608 CATCAAGAGAGATATCTCTCACTAAGCTATGACCAACCCCTTGTGCTGCTGCTGCTG 667
 887 CATGTACATGCTGATCTGCGGCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946
 668 CATGTACATGCTGATCTGCGGCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 727
 947 CACAGAAGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1006
 728 CACAGAAGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
 1007 GGCAGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
 788 GGCAGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847
 1067 CATGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126

Db 848 CACTGGACCAATGCTCTCATTCACATCAACCAAGCCCTGGTGCCCGCTACATCTT 907
 Qy 1127 CAAAGCACTCAAGTCTCTGGAATAAAGAACTCTCAGGGCTCTCGTGTATCTCT 1186
 Db 908 CAAAGCACTCAAGTCTCTGGAATAAAGAACTCTCAGGGCTCTCGTGTATCTCT 967
 Qy 1187 GGCCTCTGGCAGGGCTGCACTCAGGATACCTGCTGCTTCAGAGTGAATTCCTCAT 1246
 Db 968 GGCCTCTGGCAGGGCTGCACTCAGGATACCTGCTGCTTCAGAGTGAATTCCTCAT 1027
 Qy 1247 TGTATATGGAAGAGCAGGCTGCCAGGCTCATTCAGAGAGCCCACTTGACAGCT 1306
 Db 1028 TGTATATGGAAGAGCAGGCTGCCAGGCTCATTCAGAGAGCCCACTTGACAGCT 1087
 Qy 1307 GGCCTCATTAATGCTCTCAGGCTCTTACTATTGAGCAACAGACCATCACTGAGCT 1366
 Db 1088 GGCCTCATTAATGCTCTCAGGCTCTTACTATTGAGCAACAGACCATCACTGAGCT 1147
 Qy 1367 CTTCATGGGTAACTCCAGTACTGCTCTTGCCTCTTCACTGGGACAAATGGCTTAAAGT 1426
 Db 1148 CTTCATGGGTAACTCCAGTACTGCTCTTGCCTCTTCACTGGGACAAATGGCTTAAAGT 1207
 Qy 1427 GATATAATCATCTATTTCTCTGGCCACATCTTCTCTGAGCTTACTATTCATATTGCC 1486
 Db 1208 GATATAATCATCTATTTCTCTGGCCACATCTTCTCTGAGCTTACTATTCATATTGCC 1267
 Qy 1487 TATATATTCAGAAAGCAATGTCGCAAGAGAGAAAGTAAAGATGAATTAATCCAT 1546
 Db 1268 TATATATTCAGAAAGCAATGTCGCAAGAGAGAAAGTAAAGATGAATTAATCCAT 1327
 Qy 1547 TTCCCTGTGGCTGTGTCGGGACCTGGTGAGAAACTACTCTCTCTCTTTTCAAGCACT 1606
 Db 1328 TTCCCTGTGGCTGTGTCGGGACCTGGTGAGAAACTACTCTCTCTCTTTTCAAGCACT 1387
 Qy 1607 CCTTTGGCCCAAGAGAGAAATGGAAGAACCCAGGAGGTGAGAAAGATGCTTCCAGCT 1666
 Db 1388 CCTTTGGCCCAAGAGAGAAATGGAAGAACCCAGGAGGTGAGAAAGATGCTTCCAGCT 1447
 Qy 1667 GTGCTCTGCTGCAAGCAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTGTGGAGA 1726
 Db 1448 GTGCTCTGCTGCAAGCAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTGTGGAGA 1507
 Qy 1727 AGGCGCTTGTGCTTGTACCCACGCTGGAATGAGTGGGGATCTCAGCTCACCGCAAC 1786
 Db 1508 AGGCGCTTGTGCTTGTACCCACGCTGGAATGAGTGGGGATCTCAGCTCACCGCAAC 1567
 Qy 1787 CTCACCTCTGGGTCAAGTATTTCTGCTCTCAGCTCCCAAGTACTGGGAATACA 1846
 Db 1568 CTCACCTCTGGGTCAAGTATTTCTGCTCTCAGCTCCCAAGTACTGGGAATACA 1627
 Qy 1847 GGCAGGCCACCAATGCCAGCTAATTTTGTATTTTCAAGTGAAGAGGGATTTTCAACAGCT 1906
 Db 1628 GGCAGGCCACCAATGCCAGCTAATTTTGTATTTTCAAGTGAAGAGGGATTTTCAACAGCT 1687
 Qy 1907 TGGCCAGGCTGTGCTTGAATCTCTGACCGCAAGTATTCACCGCTCCGCTCCCAAG 1966
 Db 1688 TGGCCAGGCTGTGCTTGAATCTCTGACCGCAAGTATTCACCGCTCCGCTCCCAAG 1747
 Qy 1967 TGGCCAGGCTGTGCTTGAATCTCTGACCGCAAGTATTCACCGCTCCGCTCCCAAG 2026
 Db 1748 TGGCCAGGCTGTGCTTGAATCTCTGACCGCAAGTATTCACCGCTCCGCTCCCAAG 1807
 Qy 2027 AGCAGAGGGGCTCAATCTCCCTCTGATTTCCCTCCCAATGCAATGCTTATCTCCCA 2086
 Db 1808 AGCAGAGGGGCTCAATCTCCCTCTGATTTCCCTCCCAATGCAATGCTTATCTCCCA 1867
 Qy 2087 TGTAGCAGGAATCTATTTGTTTTTCTTCTGCAATTTACTATGATTTGTATGTGCG 2146
 Db 1868 TGTAGCAGGAATCTATTTGTTTTTCTTCTGCAATTTACTATGATTTGTATGTGCG 1927
 Qy 2147 TTAGCAGCAGCCCTCCCAATGAGGGGGGTGAGAGGGGTGCAAGGCTGCTCTCCACTT 2206
 Db 1928 TTAGCAGCAGCCCTCCCAATGAGGGGGGTGAGAGGGGTGCAAGGCTGCTCTCCACTT 1987

Qy 2207 TTTTACCTGGAACTGATTAGATAAATCACTCTGTTGTTTCAGTTTTCAAAAAA 2266
 Db 1988 TTTTACCTGGAACTGATTAGATAAATCACTCTGTTGTTTCAGTTTTCAAAAAA 2047
 Qy 2267 AAAAAA 2273
 Db 2048 AAAAAA 2054
 RESULT 5
 AX311926 1842 bp DNA linear PAT 09-JAN-2002
 LOCUS AX311926
 DEFINITION Sequence 2435 from Patent WO0194629.
 ACCESSION AX311926
 VERSION AX311926.1 GI:18122560
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 Young, P.E., Augustus, M., Carter, K.C., Edner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 TITLE
 JOURNAL
 Patent: WO 0194629-A 2435 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source
 1..1842
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 413 a 525 c 413 g 491 t
 ORIGIN
 Query Match 81.0%; Score 1842; DB 6; Length 1842;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 279 TACTCATTCACCTCTTCCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGA 338
 Db 1 TACTCATTCACCTCTTCCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGA 60
 Qy 339 AACCGCTCTACCACTCCCTGCTGTGTAATGCTTCAAGTCCATCTTGCATTAATG 398
 Db 61 AACCGCTCTACCACTCCCTGCTGTGTAATGCTTCAAGTCCATCTTGCATTAATG 120
 Qy 399 GGCAGCACATCACTGCGCTCTCACTACCTTTTCTTCAAGATGGCTTACCTTGTGCT 458
 Db 121 GGCAGCACATCACTGCGCTCTCTCACTACCTTTTCTTCAAGATGGCTTACCTTGTGCT 180
 Qy 459 GGATACATTAACCTGACCGGCACTACATATCAAGTGAAGCAATGCAATGTGTGT 518
 Db 181 GGATACATTAACCTGACCGGCACTACATATCAAGTGAAGCAATGCAATGTGTGT 240
 Qy 519 CTGACTTTGAAGCTATTTGCTTGTGCTGTGACTTCTTGAAGAGGAGGAAATACAGAT 578
 Db 241 CTGACTTTGAAGCTATTTGCTTGTGCTGTGACTTCTTGAAGAGGAGGAAATACAGAT 300
 Qy 579 TCTCTGCTCTGACCAAGAAATATATGCAATGCTGCTTCTTCCCTGCTGAAGTT 638
 Db 301 TCTCTGCTCTGACCAAGAAATATGCAATGCTGCTTCTTCCCTGCTGAAGTT 360
 Qy 639 GCTGCTTCTCTACTTCTATATGAGGCTTCTTGTGAGGCGCCAGTCTCAATGATCAC 698
 Db 361 GCTGCTTCTCTACTTCTATATGAGGCTTCTTGTGAGGCGCCAGTCTCAATGATCAC 420
 Qy 699 TACATGAAGCTGTGAGAGAGAGTGTATGACATACCAAGAAATATCAAAACGATC 758
 Db 421 TACATGAAGCTGTGAGAGAGAGTGTATGACATACCAAGAAATATCAAAACGATC 480
 Qy 759 ATTCTGCTTCAAGCGCTAGTGTGGCTTTTCTACTAGTGGGCTACACATGCTC 818

Db	481	ATTCTCTCTCTCAAGGCGCTGAGTCTG6GCGTTTTCACCTAGTGGGCTACACATGCTCTC	540
Qy	819	AGCCCCCATCAAGAAAGCTATCTCTCACTGAAAGCTATGCAACACACCCCTTTGCG	878
Db	541	AGCCCCCAATCAAGAAAGCTATCTCTCTCACTGAAAGCTATGCAACACACCCCTTTGCG	600
Qy	879	TTCCGCTGCATGTACATGCTCATCTGGGAGCAAGTTTGCTGTGTCAAATATATGCTACTGT	938
Db	601	TTCCGCTGCATGTACATGCTCATCTGGGAGCAAGTTTGCTGTGTCAAATATATGCTACTGT	660
Qy	939	TGCGTGTGCACAGAAAGAGTATGCAATTTTGAACGGGCGCTGCGCTTCAATGCTTTGAAGAA	998
Db	661	TGCGTGTGCACAGAAAGAGTATGCAATTTTGAACGGGCGCTTCAATGCTTTGAAGAA	720
Qy	999	AAGGGCAAGGCAAGTGGGATGCGCTGTGCGCAACATGAAAGGTGTGGCTTTTGAACAAAC	1058
Db	721	AAGGGCAAGGCAAGTGGGATGCGCTGTGCGCAACATGAAAGGTGTGGCTTTTGAACAAAC	780
Qy	1059	CCCCGCTTCACTGAGCACCATTTGCGCTCATTTCAACATCAACACCAAGCGCTGGGTGGCCGC	1118
Db	781	CCCCGCTTCACTGAGCACCATTTGCGCTCATTTCAACATCAACACCAAGCGCTGGGTGGCCGC	840
Qy	1119	TACATCTTCAAAACGACTCAAGTCTCTTGGAATAAAGACTCTCAAGGGTCTCTGCTTG	1178
Db	841	TACATCTTCAAAACGACTCAAGTCTCTTGGAATAAAGACTCTCAAGGGTCTCTGCTTG	900
Qy	1179	CTATTCTCGGCGCTCTGAGCAGGCGCTGACTCAGAGTACTCGTGTCTGCTTCAATGAA	1238
Db	901	CTATTCTCGGCGCTCTGAGCAGGCGCTGACTCAGAGTACTCGTGTCTGCTTCAATGAA	960
Qy	1239	TTCCCTCATTTGTATTTGGAAGAAGACAGGCTGCCAGGCTCATTTCAAGAGAGCCCACTTG	1298
Db	961	TTCCCTCATTTGTATTTGGAAGAAGACAGGCTGCCAGGCTCATTTCAAGAGAGCCCACTTG	1020
Qy	1299	AGCAAGCTGGCGCGCATTAAGTCTGCTCCAGGCGCTTACTATTTTGTGTGCAACAGACCATC	1358
Db	1021	AGCAAGCTGGCGCGCATTAAGTCTGCTCCAGGCGCTTACTATTTTGTGTGCAACAGACCATC	1080
Qy	1359	CAGTGGCTCTTCATGAGGTTAAGTCTGCAAGTACTGCTCTTGCCTTTCACGCTGGGACAAATGG	1418
Db	1081	CAGTGGCTCTTCATGAGGTTAAGTCTGCAAGTACTGCTCTTGCCTTTCACGCTGGGACAAATGG	1140
Qy	1419	CTTAAGGTGTATTAATTCATCTATTTCTTGGCCACATCTTCTTCTGAGGCTACTATTC	1478
Db	1141	CTTAAGGTGTATTAATTCATCTATTTCTTGGCCACATCTTCTTCTGAGGCTACTATTC	1200
Qy	1479	ATATTGCTTTATATTCACAAAGCAATGCTGCGCAAGAAAGAAAGATTAAAGAAATGGA	1538
Db	1201	ATATTGCTTTATATTCACAAAGCAATGCTGCGCAAGAAAGAAAGATTAAAGAAATGGA	1260
Qy	1539	TAAATCAATTTCCCTGTGTGCGCTGTGTGGGAACTGTGTCAAGAACTATCTGCTCTCCCTTTC	1598
Db	1261	TAAATCAATTTCCCTGTGTGCGCTGTGTGGGAACTGTGTCAAGAACTATCTGCTCTCCCTTTC	1320
Qy	1599	ACAGCACTCTTTGCGCGCCAGAGACAGAGATGGAAGAACACAGGAGGTGGAAGATCGATGC	1658
Db	1321	ACAGCACTCTTTGCGCGCCAGAGACAGAGATGGAAGAACACAGGAGGTGGAAGATCGATGC	1380
Qy	1659	TTCCAGCTGTGCTGTGCTGCGACGCAAGTCTTCAATTTGGGCGCAAGGGGAAACTTTT	1718
Db	1381	TTCCAGCTGTGCTGTGCTGCGACGCAAGTCTTCAATTTGGGCGCAAGGGGAAACTTTT	1440
Qy	1719	TTTGGAGAAAGGGCTTGTCTTGTATCCCAAGCTGTGAATGCAAGTGGCGGAACTCAAGCTC	1778
Db	1441	TTTGGAGAAAGGGCTTGTCTTGTATCCCAAGCTGTGAATGCAAGTGGCGGAACTCAAGCTC	1500
Qy	1779	ACCGGAACCTCAACCTCTCGGGTTCAAGTGAATTTCTGAGCTCAAGCTCCCAAGTATGCTG	1838
Db	1501	ACCGGAACCTCAACCTCTCGGGTTCAAGTGAATTTCTGAGCTCAAGCTCCCAAGTATGCTG	1560
Qy	1839	GGAAATACAGCAGCGCACCATGCGCAGCTAATTTTGTATTTTCAATGAAAGGGAGATT	1898

QY	Db	1561	GGAAATPAGGACAGCCACCCATGGCCACGTAATTTTGTATTTCAGTAGAAAGCGGATTT	1628
QY	1899	CACCACGTTG3CCAGGCTGTGTCTTCGAATCTCTTGAACCGCAAGTATCCACCCGCTCCGCC	1958	
Db	1621	CACCACGTTGGCCAGGCTGTGTCTTCGAATCTCTTGAACCGCAAGTATCCACCCGCTCCGCC	1680	
QY	1959	TCCCAAGTCTGGGATTTTACAGGCGTAGAGCACCGTCCGGGCCCAAGGGGAATCTTT	2018	
Db	1681	TCCCAAGTCTGGGATTTTACAGGCGTAGAGCACCGTCCGGGCCCAAGGGGAATCTTT	1740	
QY	2019	GTGGAGGAGACAGAGGGGCTCACATCTCCCTCGATTTCCTCCCATGACATTTGCTTATC	2078	
Db	1741	GTGGAGGAGACAGAGGGGCTCACATCTCCCTCGATTTCCTCCCATGACATTTGCTTATC	1800	
QY	2079	TCTCCCATCTAGCCAGCAATCTATTGTGTTTTTCTTCTGCC	2120	
Db	1801	TCTCCCATCTAGCCAGCAATCTATTGTGTTTTTCTTCTGCC	1842	
RESULT 6				
AX410769	AX410769	1842 bp	DNA	linear
LOCUS	Sequence 3416 from Patent WO0229103.			
DEFINITION	AX410769			
ACCESSION	AX410769.1	GI:21443474		
VERSION				
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	Alvares,C., Horne,D., Pears-da-Silva,S. and Vockley,J.G.			
JOURNAL	Gene expression profiles in liver cancer			
GENE LOGIC INC (US)	Patent: WO 0229103-A 3416 11-APR-2002;			
FEATURES				
Source	location/Qualifiers			
	1..1842			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
	/note="EMBL/GenBank Accession No. U72515"			
BASE COUNT	413 a 525 c 413 g 491 t			
ORIGIN				
Query Match	81.0%; Score 1842; DB 6; Length 1842;			
Best Local Similarity	100.0%; Pred. No. 0;			
Matches 1842; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	2719	TACCTCATCCACCTCTTCCATACCTTTACAGGCTCTTCATATTCGTTATTTTAACTTTGA	338	
Db	1	TACCTCATCCACCTCTTCCATACCTTTACAGGCTCTTCATATTCGTTATTTTAACTTTGA	60	
QY	339	AACCAAGCTCTACACTCCCTGTGTGATGTGCTTCACTTCATCTCGACTAATG	398	
Db	61	AACCAAGCTCTACACTCCCTGTGTGATGTGCTTCACTTCATCTCGACTAATG	120	
QY	399	GGCCGACCATCACTGCGGCTCTCACTATCTTTGCTTCCAGATGGCTACCTTGGCT	458	
Db	121	GGCCGACCATCACTGCGGCTCTCACTATCTTTGCTTCCAGATGGCTACCTTGGCT	180	
QY	459	GGATTAATTAACCTGCGACCGGCAACTAGATATCAAGTGAACAATGCCACATTTGCT	518	
Db	181	GGATTAATTAACCTGCGACCGGCAACTAGATATCAAGTGAACAATGCCACATTTGCT	240	
QY	519	CTGACTTTGAAGCTGATGTTGTTGGCTTTGACTTATTTGACGAGGAGAAAGATCAGAT	578	
Db	241	CTGACTTTGAAGCTGATGTTGTTGGCTTTGACTTATTTGACGAGGAGAAAGATCAGAT	300	
QY	579	TCTTTGCTCTGAGCAACAGAAATATGCCATACGTGTGTCTTCTCTGCTGGAAGTT	638	
Db	301	TCTTTGCTCTGAGCAACAGAAATATGCCATACGTGTGTCTTCTCTGCTGGAAGTT	360	
QY	639	GCTGTTTCTCTTACTTCTATGGGGCTTTTGTGATGGGCCCCAGTTCTCATATGATATAC	698	

```

|||||
Db 361 GGTGGTTTCTCCACTTCTATGGGGGCTTCTGGTGGGGCCCGAGTTCTCAATGAATCAC 420
|||
Qy 699 TACATGAAGCTGTGAGGAGAGAGCTGATGACATACCGAGAAAGATACCAGAACGATC 758
|||
Db 421 TACATGAAGCTGTGAGGAGAGAGCTGATGACATACCGAGAAAGATACCAGAACGATC 480
|||
Qy 759 ATTCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACCTAGTGGGCTACACACTGTC 818
|||
Db 481 ATTCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACCTAGTGGGCTACACACTGTC 540
|||
Qy 819 AGCCCCACATCAAGAGAGATATCTCTCACTGAAGACTATGACAAACCCCTTTCTG 878
|||
Db 541 AGCCCCACATCAAGAGAGATATCTCTCACTGAAGACTATGACAAACCCCTTTCTG 600
|||
Qy 879 TTCCGCTGATGACATGCTGATCTGGGGCAAGTTTGTCTGACAAATATGCACTGT 938
|||
Db 601 TTCCGCTGATGACATGCTGATCTGGGGCAAGTTTGTCTGACAAATATGCACTGT 660
|||
Qy 939 TGGCTGTGACAGAGAGATATGATCTTTGACGGGCTTGCGCTTCAATGGCTTTGAGAA 998
|||
Db 661 TGGCTGTGACAGAGAGATATGATCTTTGACGGGCTTGCGCTTCAATGGCTTTGAGAA 720
|||
Qy 999 AAGGCGAAGCAAGTGGAGTGTCTGTGCAACATGAGGTGTGCTCTTTGAAACAAAC 1058
|||
Db 721 AAGGCGAAGCAAGTGGAGTGTCTGTGCAACATGAGGTGTGCTCTTTGAAACAAAC 780
|||
Qy 1059 CCCCCTTCACTGGGCAACATGCTGATTCATCAATCAACCAACCCCTGGTGGCCCGC 1118
|||
Db 781 CCCCCTTCACTGGGCAACATGCTGATTCATCAATCAACCAACCCCTGGTGGCCCGC 840
|||
Qy 1119 TACATCTCAAGACTCAAGTCTCTGGAATAAAGAACTCTGAGGATCTCTGTTG 1178
|||
Db 841 TACATCTCAAGACTCAAGTCTCTGGAATAAAGAACTCTGAGGATCTCTGTTG 900
|||
Qy 1179 CTATTCCTGGCCCTCTGGGACGGGCTGCACTCAGATACCTGGTCTGCTTCAGATGAA 1238
|||
Db 901 CTATTCCTGGCCCTCTGGGACGGGCTGCACTCAGATACCTGGTCTGCTTCAGATGAA 960
|||
Qy 1239 TTCCCTATGTTATTTGTGGAAGAGAGGCTGCCAGGCTCATTCAGAGAGCCCACTG 1298
|||
Db 961 TTCCCTATGTTATTTGTGGAAGAGAGGCTGCCAGGCTCATTCAGAGAGCCCACTG 1020
|||
Qy 1299 ACCAAGCTGGCGCCATCTACTGCTCTCCAGCCCTTCTACTATTTGTGGAACAGACATC 1358
|||
Db 1021 ACCAAGCTGGCGCCATCTACTGCTCTCCAGCCCTTCTACTATTTGTGGAACAGACATC 1080
|||
Qy 1359 CACTGCTCTTCAATGGGTTACTTCATGACTGCTTCTGCTCTTCACTGGAACAAATG 1418
|||
Db 1081 CACTGCTCTTCAATGGGTTACTTCATGACTGCTTCTGCTCTTCACTGGAACAAATG 1140
|||
Qy 1419 CTTAAGGTATTAATTCATCTATTTCTTGGGCAATCTTCTCTGAGCTACTATTC 1478
|||
Db 1141 CTTAAGGTATTAATTCATCTATTTCTTGGGCAATCTTCTCTGAGCTACTATTC 1200
|||
Qy 1479 ATATTGCTTATATTCAGAAAGCATGTGTGCCAAGAAAGAGAGTTAAGAAAGATGAA 1538
|||
Db 1201 ATATTGCTTATATTCAGAAAGCATGTGTGCCAAGAAAGAGAGTTAAGAAAGATGAA 1260
|||
Qy 1539 TAATCATTTTCCCTGTGGCTGTGGGAGCTGTGACAGAACTACTGCTCTCCCTTTTC 1598
|||
Db 1261 TAATCATTTTCCCTGTGGCTGTGGGAGCTGTGACAGAACTACTGCTCTCCCTTTTC 1320
|||
Qy 1599 ACAGCATCTCTTGGCCCCAGAGAGAGATGAGAAAGCCAGGAGGTGGAATGCATGC 1658
|||
Db 1321 ACAGCATCTCTTGGCCCCAGAGAGAGATGAGAAAGCCAGGAGGTGGAATGCATGC 1380
|||
Qy 1659 TTCCAGCTGTGCTGTGCTGCGAGCAAGTCTTCAATTTGGGGCAAGGGGAACTTTT 1718
|||
Db 1381 TTCCAGCTGTGCTGTGCTGCGAGCAAGTCTTCAATTTGGGGCAAGGGGAACTTTT 1440
|||
Qy 1719 TTGGAGAAAGGCGTCTTGTCTTGTCAACGCTGGAATGCACTGGCGGATCTCAGCTC 1778
|||

```

```

Db 1441 TTGGAGAAAGCGTCTTCTTGTTCACCCAGCTGGAATGAGTGGCGGATCTCAGCTC 1500
|||
Qy 1779 ACCGAACTCCACTCCTGGGTTCAAGTATTTTCCGCTCAGGCTCCCAAGTAGCTG 1838
|||
Db 1501 ACCGAACTCCACTCCTGGGTTCAAGTATTTTCCGCTCAGGCTCCCAAGTAGCTG 1560
|||
Qy 1839 GGAATACAGGACGCGCACCATGCGCAGCTAAATTTTGTATTTTCAGTAAACGGGATTT 1898
|||
Db 1561 GGAATACAGGACGCGCACCATGCGCAGCTAAATTTTGTATTTTCAGTAAACGGGATTT 1620
|||
Qy 1899 CACCAAGTGGCGAGGCTGTCTGCACTCTGACCGCAAGTATCCACCGCTCCGCC 1958
|||
Db 1621 CACCAAGTGGCGAGGCTGTCTGCACTCTGACCGCAAGTATCCACCGCTCCGCC 1680
|||
Qy 1959 TCCCAAGTGGCGAGGCTGTCTGCACTCTGACCGCAAGTATCCACCGCTCCGCC 2018
|||
Db 1681 TCCCAAGTGGCGAGGCTGTCTGCACTCTGACCGCAAGTATCCACCGCTCCGCC 1740
|||
Qy 2019 GTGGAGAGAGAGAGGCTCACAATCTCCCTCTGATTCGCCCATGACATTTGCTTATC 2078
|||
Db 1741 GTGGAGAGAGAGAGGCTCACAATCTCCCTCTGATTCGCCCATGACATTTGCTTATC 1800
|||
Qy 2079 TCTCCCATCTAGCCAGGAATCTATGTGTTTCTTCTGCC 2120
|||
Db 1801 TCTCCCATCTAGCCAGGAATCTATGTGTTTCTTCTGCC 1842
|||
RESULT 7
HSU72515 1842 bp mRNA linear PRI 24-JUL-1997
LOCUS HSU72515
DEFINITION Human C3f mRNA, complete cds.
ACCESSION U72515
VERSION U72515.1 GI:1673519
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Ansel-Lari, M.A., Shen, Y., Muzny, D.M., Lee, W. and Gibbs, R.A.
TITLE Large-scale sequencing in human chromosome 12p13: experimental and
JOURNAL computational gene structure determination
GENEID 97228904
MEDLINE 9074930
PUBMED 2 (bases 1 to 1842)
REFERENCE
AUTHORS Ansel-Lari, M.A., Shen, Y., Muzny, D.M., Lee, W. and Gibbs, R.A.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1996) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
FEATURES
source location/Qualifiers
1..1842
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13"
1..1842
/gene="C3f"
1..1842
/gene="C3f"
118..1263
/gene="C3f"
/note="Similar to ESTs with GenBank Accession Numbers
H45806 and H45837; similar to S. cerevisiae ORF YOR175c
encoded by GenBank Accession Number Z75083; see
corresponding genomic sequence in GenBank Accession Number
U72506"
/codon_start=1
/product="C3f"
/protein_id="AAC51640.1"
/db_xref="GI:1673520"
/translation="MGRTITAVLTFQFMAYLLAGYVYATAGNYDIKMTBPHCVLTL

```

KLIGLADVYEDGKDNLSSEBQKXAIRBPSLLBVAESRYFYGAFLVIGPORSNMHY
MKLVGBELIDIPKIPNSIIPALRSLISLGFVIVGTLLPTEBYLITBEDNDHP
WERCMYMLWIKFVLYKYVTCMLVEGVCLLIGLNGEKEKAKMADKANNKWL
ETNRPETIASFININAMVARYIFRKLFLGNKSLSQSLFLTALHMGSLGTLV
CFQMEFLIVVEROARLIOESPTELSKLAIVLOPFPYVLOOTIHLFMXGSMATFC
LFTMDKMLKVKYSIYFLIGHIFLSLFIPIYIKHAWPREKXKME"

BASE COUNT 413 a 525 c 413 g 491 t

ORIGIN

Query Match 81.0%; Score 1842; DB 9; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 TACCTCATCACTCTTCCATACCTTTACAGGCTCTCAATGGTATTTTAACTTTGA 338
DB 1 TACCTCATCACTCTTCCATACCTTTACAGGCTCTCAATGGTATTTTAACTTTGA 60
QY 339 AACAGCTCAACCACTCCCTGCTGTGATTTGCTGATTCCTCACTCTTCACTAATG 398
DB 61 AACAGCTCAACCACTCCCTGCTGTGATTTGCTGATTCCTCACTCTTCACTAATG 120
QY 399 GGGCGACATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTTACCTTGGCT 458
DB 121 GGGCGACATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTTACCTTGGCT 180
QY 459 GGAATCAATTAACATGCGCACTGCGCACTACGATATCAAGTGGACATGCTTGT 518
DB 181 GGAATCAATTAACATGCGCACTGCGCACTACGATATCAAGTGGACATGCTTGT 240
QY 519 CTACCTTTGAAGCTGATTTGCTGCTGTGATCTACTTTGACGAGGAGAAAGATCAAT 578
DB 241 CTACCTTTGAAGCTGATTTGCTGCTGTGATCTACTTTGACGAGGAGAAAGATCAAT 300
QY 579 TCCCTGTCTCTGAGCAAGAAATATGCCATACGTGTGCTCTTCCCTGCTGGAAT 638
DB 301 TCCCTGTCTCTGAGCAAGAAATATGCCATACGTGTGCTCTTCCCTGCTGGAAT 360
QY 639 GCTGTGCTTCTCTAATCTTCTATGCGGCTTCTTGTGAGGCGCCCACTTCAATGATAC 658
DB 361 GCTGTGCTTCTCTAATCTTCTATGCGGCTTCTTGTGAGGCGCCCACTTCAATGATAC 420
QY 699 TACATGAAGCTGTGAGGAGAGCTGATTTGACATACGAGAAAGATACCAACAGCATC 758
DB 421 TACATGAAGCTGTGAGGAGAGCTGATTTGACATACGAGAAAGATACCAACAGCATC 480
QY 759 ATTCTGCTCTCAAGCGCTGAGCTGTGGGCTTTTCTACTAGTGGCTACACATGCTC 818
DB 481 ATTCTGCTCTCAAGCGCTGAGCTGTGGGCTTTTCTACTAGTGGCTACACATGCTC 540
QY 819 AGCCCCCAATCAAGAAAGTATCTCTCACTGAAACATATGACACACCCCTTGTG 878
DB 541 AGCCCCCAATCAAGAAAGTATCTCTCACTGAAACATATGACACACCCCTTGTG 600
QY 879 TTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 938
DB 601 TTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 939 TGGCTGTGACAGAAAGATGATGATTTTGAACGCGCTGTGCTTCAATGCTTTGAAGAA 998
DB 661 TGGCTGTGACAGAAAGATGATGATTTTGAACGCGCTGTGCTTCAATGCTTTGAAGAA 720
QY 999 AAGGGCAAGGCAAGTGGATGCTGTGCGCAACATGAAGTGTGGCTTTTGAAGAAAC 1058
DB 721 AAGGGCAAGGCAAGTGGATGCTGTGCGCAACATGAAGTGTGGCTTTTGAAGAAAC 780
QY 1059 CCGCGCTTCACTGACCATTTGCTCATTTCAATCAACACCAACGCGCTGGGTGGCCGC 1118
DB 781 CCGCGCTTCACTGACCATTTGCTCATTTCAATCAACACCAACGCGCTGGGTGGCCGC 840
QY 1119 TACATCTTCAACACGACTCAAGTTCTTGAAGTAAAGAACTCTCAAGGCTCTCTGTT 1178
DB 841 TACATCTTCAACACGACTCAAGTTCTTGAAGTAAAGAACTCTCAAGGCTCTCTGTT 900

QY 1179 CTATTCCTGGCCCTCTGAGACGGCCCTGACCTGACGATACCTGCTGCTTCCAGATGAA 1238
DB 901 CTATTCCTGGCCCTCTGAGACGGCCCTGACCTGACGATACCTGCTGCTTCCAGATGAA 960
QY 1239 TTCCTCATGTTATTTGTGAAAGACAGGCTGCGAGGCTCATTTCAAGAGAGCCCACTG 1298
DB 961 TTCCTCATGTTATTTGTGAAAGACAGGCTGCGAGGCTCATTTCAAGAGAGCCCACTG 1020
QY 1299 AGCAAGCTGGCCGCAATTAATCTGCTCTGACGCTTCTACTATTTGGTGCACAGACATC 1358
DB 1021 AGCAAGCTGGCCGCAATTAATCTGCTCTGACGCTTCTACTATTTGGTGCACAGACATC 1080
QY 1359 CACTGCTCTTCAATGAGGCTTACCTCAATGATGCTGCTGCTTCACTGAGGAGCAATATG 1418
DB 1081 CACTGCTCTTCAATGAGGCTTACCTCAATGATGCTGCTGCTTCACTGAGGAGCAATATG 1140
QY 1419 CTTAAGGTATTAATTCATATTTCTTGGCCACATCTTCTCTGAGCTTACTATTC 1478
DB 1141 CTTAAGGTATTAATTCATATTTCTTGGCCACATCTTCTCTGAGCTTACTATTC 1200
QY 1479 ATATGCTTATTAATTCACAAAGCAATGTTGCCAAGAAAGAAAGTAAAGATGAA 1538
DB 1201 ATATGCTTATTAATTCACAAAGCAATGTTGCCAAGAAAGAAAGTAAAGATGAA 1260
QY 1539 TAATCAATTTCCCTGCTGCTGCTGCGAGCTGCTGCAAAACTACTGCTTCCCTTTC 1598
DB 1261 TAATCAATTTCCCTGCTGCTGCTGCGAGCTGCTGCAAAACTACTGCTTCCCTTTC 1320
QY 1599 ACAGCACTCTTTCGCCAGAGAGAGATGAAAGCCAGGAGGTGAAAGATTCATG 1658
DB 1321 ACAGCACTCTTTCGCCAGAGAGAGATGAAAGCCAGGAGGTGAAAGATTCATG 1380
QY 1659 TTCAGCTGTGCTGCTGCTGCGAGCAAGCTTCTTATTTGGGCGCAAGGGAAACTTTT 1718
DB 1381 TTCAGCTGTGCTGCTGCTGCGAGCAAGCTTCTTATTTGGGCGCAAGGGAAACTTTT 1440
QY 1719 TTTGAGAAAGGCTCTTGTGCTGCTGCGAGCAAGCTGGAAGTGGCGGGATCTGAGCT 1778
DB 1441 TTTGAGAAAGGCTCTTGTGCTGCTGCGAGCAAGCTGGAAGTGGCGGGATCTGAGCT 1500
QY 1779 ACCGCAACTCTCACTCTGAGGCTTCAAGTATTTCTGCTCAGCTCCCAAGTATG 1838
DB 1501 ACCGCAACTCTCACTCTGAGGCTTCAAGTATTTCTGCTCAGCTCCCAAGTATG 1560
QY 1839 GGAATACAGGACGCCACCATGCTCAATTTTGTATTTTCAATGAAAGGGAAT 1898
DB 1561 GGAATACAGGACGCCACCATGCTCAATTTTGTATTTTCAATGAAAGGGAAT 1620
QY 1899 CACCAAGTTGGCCAGGCTGCTGCAACTCTGACCGCAAGTATCACCCGCTCGGC 1958
DB 1621 CACCAAGTTGGCCAGGCTGCTGCAACTCTGACCGCAAGTATCACCCGCTCGGC 1680
QY 1959 TCCCAAGTGTGGATTTACAGGCTGAGCACTGCGCGCCCAAGGGGAACTCT 2018
DB 1681 TCCCAAGTGTGGATTTACAGGCTGAGCACTGCGCGCCCAAGGGGAACTCT 1740
QY 2019 GTGGAGAGAGAGAGGCTCACATCTCCCTCTGATTTCCCATGACATTCCTTATC 2078
DB 1741 GTGGAGAGAGAGAGGCTCACATCTCCCTCTGATTTCCCATGACATTCCTTATC 1800
QY 2079 TCTCCCATCTAGCCAGGAATCATTTGTTTCTTCTGCC 2120
DB 1801 TCTCCCATCTAGCCAGGAATCATTTGTTTCTTCTGCC 1842

RESULT 8
LOCUS AX073560
DEFINITION Sequence 12 from Patent WO0104297.
ACCESSION AX073560
VERSION AX073560.1 GI:12709973
KEYWORDS
SOURCE Homo sapiens (human)

JOURNAL

Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES

Location/Qualifiers
1..1573

Source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TST00701"
/cissue_type="testis"
/clone_id="TST"
/note="Cloning vector: pME18SFL3"

BASE COUNT 358 a 436 c 378 g 401 t

ORIGIN

Query Match 61.1%; Score 1388.2; DB 9; Length 1573;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1408; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

814 TGCTGACCCCAATCAGAGAGATCTCTCTGAGAACTATGACACACCCCT 873
133 TGTTCATCAGCCGACGATGAAACCTTCGGGGCAGGGCTGTGCAACACCCCT 192
874 TCTGTCCTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
193 TCTGTCCTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 252
934 CCGTGGCTGTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 993
253 CCGTGGCTGTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
994 AAGAAAGGCGCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053
313 AAGAAAGGCGCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
1054 CAAACCCCGCTTCACTGAGCACTTGCCTTCACTGAGCACTTGCCTTCACTGAGCA 1113
373 CAAACCCCGCTTCACTGAGCACTTGCCTTCACTGAGCACTTGCCTTCACTGAGCA 432
1114 CCGGCTGATCTTCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
433 CCGGCTGATCTTCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
1174 CGTTGCTATTCCTGAGCGCTTGCAGCGGCTGCACTGAGATACCTGCTGCTTCAGA 1233
493 CGTTGCTATTCCTGAGCGCTTGCAGCGGCTGCACTGAGATACCTGCTGCTTCAGA 552
1234 TGGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293
553 TGGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612
1294 CCGGAGCAAGCTGCGCGGCTTCACTGAGCGGCTTCACTGATGATGATGATGATGAT 1353
613 CCGGAGCAAGCTGCGCGGCTTCACTGAGCGGCTTCACTGATGATGATGATGATGAT 672
1354 CCATCCATGAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1413
673 CCATCCATGAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
1414 AATGCTTAAGGTATTAATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1473
733 AATGCTTAAGGTATTAATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 792
1474 TATTCATATTCCTTATATTCACAAAGCAATGATGAGCAAGAGAGAGAGAGAGAGAG 1533

Db 793 TATTCATATTCCTTATATTCACAAAGCAATGATGAGCAAGAGAGAGAGAGAGAGAG 852
1534 TGGATTAATCACTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1593
853 TGGATTAATCACTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 912
1594 TTTTCACAGACTCTCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1653
913 TTTTCACAGACTCTCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
1654 GATGCTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1713
973 GATGCTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
1714 TTTTTCAG 1773
1033 TTTTTCAG 1092
1774 AGCTCAGCGCACTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1833
1093 AGCTCAGCGCACTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152
1834 AGCTGGAATPACAGGAG 1893
1153 AGCTGGAATPACAGGAG 1212
1894 GATTTACACAGCTTGGCAG 1953
1213 GATTTACACAGCTTGGCAG 1272
1954 CCGCTCCCAAG 2013
1273 CCGCTCCCAAG 1332
2014 CTTTGGAG 2073
1333 CTTTGGAG 1392
2074 TTAATCTCTCCCATCTGAG 2133
1393 TTAATCTCTCCCATCTGAG 1452
2134 TGTGATGAG 2193
1453 TGTGATGAG 1512
2194 GCCTGCTCACTTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2253
1513 GCCTGCTCACTTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
2254 T 2254
1573 T 1573

RESULT 10

AK098090

LOCUS

DEFINITION

AK098090

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

2128 bp mRNA linear ROD 25-SEP-2002
Mus musculus cDNA fls, clone TRACH2004292, highly similar to Human
c3f mRNA.
AK098090.1 GI:21758026
Oligo capping, fls (full insert sequence).
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Mutsaers, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,
Isono, Y., Kawai-Hito, Y., Saito, K., Nishikawa, T., Kimura, K.,

Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Magatsuna, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Negahari, K., Masuno, Y., Negai, K. and Isogai, T.
NEDO cDNA sequencing project
Unpublished
2 (bases 1 to 2128)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genom@seq.hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center; National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2128
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="TRACH2004292"
/note="cloning vector: pME18FLJ"
BASE COUNT 476 a 601 c 500 g 551 t
ORIGIN
Query Match 54.6%; Score 1240.4; DB 10; Length 2128;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 133; Conservative 0; Mismatches 241; Indels 18; Gaps 3;
1 GGGGGTGAAGCATACGTTTGGCCGATTCGGGGCGCGGAGCTGGGGGGTCCCTGTG 60
3 GGGGTGAAGCCACCGCGTCTGCCCAATTCGGGGCGCGGAG---CTGGGGTCCCGGG 59
61 GGGCTCCGGAGTAAAGATGGGCTCTCAGCGGAGGGGAGCGAGGGGAGTGTGGGCGC 120
60 GGGCGCCGCAATTAAGATGGGCTTACAGCGGAGCGGAGCAATGGAGAGAGCGCTGGAGC 119
121 TGGCGGGGTTTGCAGTCGGGTTTTCAGAGCTGAGCTTAAAGTTGGGAGAGTCCC 180
120 AGATCGGGGGCTGTGGCGGGGTGCGAGATCTGAGCTTAAAGTTGGGAGAGTCTC 179
181 TGGGGCGGTTCAGAACAGCGCTGCGCTGATCTTCATCTTCTGGGTTAACCCCTTTG 240
180 TGGGGCGGTTCAGAACAGCGCTGCGCTGATCTTCATCTTCTGGGTTAACCCCTTTG 239
241 CTTGTTTATGGGATTAACCTTTCTAAGAGAGCTCACTCCACTCCCTTCCANA 300
240 CTCGTTTATCCGGATTAACCTTTCTAAGAGAGCTCACTCCACTCCCTTCCANA 299
301 CTTTACAGGCTCTCAATTGCTTAATTTTAACTTTGGAACAGAGCTTACACTCCCTG 360
300 CTTACAGGCTCTCAATTGCTTAATTTTAACTTTGGAACAGAGCTTACACTCCCTG 359
361 TGTGATGTTGCTTCAAGTTCCTATCTTGCAGTAAATGGGCGGACACTGCGCTTC 420
360 TATGTTGTTGCTTCAAGTTCCTATCTTGCAGTAAATGGGCGGACACTGCGCTTC 419
421 TCACATACCTTTGCTTCCAGATGGGCTACCTCTGGCTGATTAATTAACAATGACCG 480
420 TTACATACCTTTGCTTCCAGATGGGCTACCTCTGGCTGATTAATTAACAATGACCG 479
481 GCAACTAGATATCAAGTGGACATGCCAATTTGTTCTGACTTTGAAGCTGATGTT 540
480 GTGACTAGATATCAAGTGGACATGCCAATTTGTTCTGACTTTGAAGCTGATGTT 539
541 TGGCTGTTGACTTACTTTGACGAGGGAAGATCAAAATTCCTTGTCTCTGACACAGA 600
540 TGTGATGTTGACTTACTTCAATGAGGCAAGAGCGGAAATTCCTTGTCTCTGACACAGA 599

601 AATATGCATACGTTGTTCTTCCCTGCTGAGAGTTCGTTCTTCTTACTTCTATG 660
600 AATATGCATACGTTGTTCTTCCCTGCTGAGAGTTCGTTCTTCTTACTTCTATG 659
661 GGGCTCTTCTGTTAGGGCCCGAGTTCTCAATGATCACTACATGAGCTGTGACGAG 720
660 GAGCTTCTTCTGTTAGGGCCCGAGTTCTCAATGATCACTACATGAGCTGTGACGAG 719
721 AGCTGATTAACATACAGGAAAGATACCAACAGCATATTCCTGCTTCAAGCCCTGA 780
720 AGCTGATTAACATACAGGAAAGATACCAACAGCATATTCCTGCTTCAAGCCCTGA 779
781 GTCCTGGGCTTTTCTACTACTGAGCTTACAGCTGCTACGCCCCCACTACAGAACT 840
780 GTCCTGGGCTTTTCTACTACTGAGCTTACAGCTGCTACGCCCCCACTACAGAACT 839
841 ATCTCTCACTGAAGACTATGACAAACACCCCTTGTGTTCCGCTGATGATGATGTA 900
840 ATCTCTCACTGAAGACTATGATTAACCGCCCTTGTGTTCCGCTGATGATGATGTA 899
901 TCTGGGCAAGTTGTGCTGTACAAATATGTCACCTGTTGCTGTACAGAGAGATAT 960
900 TCTGGGCAAGTTGTGCTGTACAAATATGTCACCTGTTGCTGTACAGAGAGATAT 959
961 GCATTTTGAAGGCGCTGGGCTTCAATGGCTTTGAAGAAAGGCAAGGCAAGTGGATG 1020
960 GCATTTTGAAGGCGCTGGGCTTCAATGGCTTTGAAGAAAGGCAAGGCAAGTGGATG 1019
1021 CCTGTGCCAATGAGGTGTGCTCTTTGAAGAAACCCCGCTTCACTGGACCATTTG 1080
1020 CCTGTGCCAATGAGGTGTGCTCTTTGAAGAAACCCCGCTTCACTGGACCATTTG 1079
1081 CCTATTCAATCAACCAACCGCTGGTGGCCGCTCACTCTTCAACAGACTCAAGT 1140
1080 CCTATTCAATCAACCAACCGCTGGTGGCCGCTCACTCTTCAACAGACTCAAGT 1139
1141 TCTTTGGAATTAAGAACTCTCAAGGCTCTCTGTTCTATCTTGGCCCTCTGGACAG 1200
1140 TCTTTGGAATTAAGAACTCTCAAGGCTCTCTGTTCTATCTTGGCCCTCTGGACAG 1199
1201 GCTTGCACATGAGATACCTGTGCTCTTCCAGATGGAATTCCTCATTTGTTGAAA 1260
1200 GCTTGCACATGAGATACCTGTGCTCTTCCAGATGGAATTCCTCATTTGTTGAAA 1259
1261 GACAGCTGCCAGCTCAATTAAAGAGCCCCCAGCTGAGCAAGCTGCCCCCATTAATG 1320
1260 GACAGCTGCCAGCTCAATTAAAGAGCCCCCAGCTGAGCAAGCTGCCCCCATTAATG 1319
1321 TCTTCCAGCCCTTCTACTAATTTGTCGCAACAGACATCACTGGCTCTTCAATGGTTACT 1380
1320 TCTTCCAGCCCTTCTACTAATTTGTCGCAACAGACATCACTGGCTCTTCAATGGTTACT 1379
1381 CCATGACTGCTTCTGCTCTTCAAGTGGGCAATATGCTTAAGGTATTAATCATCT 1440
1380 CCATGACTGCTTCTGCTCTTCAAGTGGGCAATATGCTTAAGGTATTAATCATCT 1439
1441 ATTTCCTTGGCAATCTTCTTCTGAGCTTACTAATTAATGCTTAAATTCAGAAAG 1500
1440 ATTTCCTTGGCAATCTTCTTCTGAGCTTACTAATTAATGCTTAAATTCAGAAAG 1499
1501 CAATGTTGCAAGGAAAGAGATTAAGAAAGATGGAATTAATC-----CATTTCCCTG 1555
1500 CAATGTTGCAAGGAAAGAGATTAAGAAAGATGGAATTAATC-----CATTTCCCTG 1554
1556 GGCCTGTGCGGAGCTGTGACAGAAATCTGCTCTTCTTCAAGAGACTCTTGTGCCC 1615
1560 GGCCTGTGCGGAGCTGTGACAGAAATCTGCTCTTCTTCAAGAGACTCTTGTGCCC 1619
1616 CAGAGCAAGATGAGAAAGCCAGGAGGTGAAAGATGATGCTTCCAGCTGTGCTTGT 1675
1620 -----AGCAGAGCCGACAGAGCTGAGATCAAGCACTCCAGCTGTGCTTGT 1669

QY 1676 CTCGACGCCAAG 1687
| | | | |
Db 1670 CGACGCCAAG 1681

RESULT 11
BC006753

LOCUS
DEFINITION
BC006753 Mus musculus gene rich cluster, C3f gene, mRNA (CDNA clone
MGC:11670 IMAGE:3709076), complete cds.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
MGC.
Mus musculus (house mouse)

REFERENCE
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Strapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Adams, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Guarnatone, P.H., Richards, S.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shvachenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Strausberg, R.
Direct Submission
Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Guarnatone, P.H., Garcia, A.M., Lu, X., Huiyik, S.W., Louisedge, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navevati,
A.N., Gibbs, R.A.

FEATURES
SOURCE
1..1898
/organism="Mus musculus"
/mol_type="mRNA"

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 17 Row: h Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarly but not
identity to protein.
Location/Qualifiers
1..1898

/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:11670 IMAGE:3709076"
/issue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_1lb="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
1..1898
/gene="Grcc3f"
/note="synonyms: C3f, PTC"
/db_xref="LocusID:14792"
/db_xref="MGI:1315211"
34..1497
/codon_start=1
/product="gene rich cluster, C3f gene"
/protein_id="AAH06753.2"
/db_xref="GI:14548370"
/db_xref="LocusID:14792"
/translation="MASTADGDMGTLEBGMRLMPGVEDLSINKLATSIGASEOALRL
ITSTPLPLAFYRHYLFYKDSYIHLPHFTTGISIAFNFGHOFHSLCYVLOFL
ILRLGRVITAVITTLCEQAYLLAGYTTAGDVDKWMPHCVTLKLIGLCTIDY
DEGKXNSLTSIQKXAIKRGVSLLEVAFSYFAFLVPOFSNMHWKLVKVGQILD
IGKPNSTIPALKEKSLGLVYLVGYTLISPHITDYLTEDYDNRPFRCVMYLM
GKPVLYKVTWMLVTEGVCLISLGFNGDEGTVRMDACAMKWLFTETPRNGTI
AEPNINWAWARVYKPKLKGKNEISGLISLFLALMGHLSGYLICFOMELFVLI
YKIVSLIRDSPLASLSLSTALDPLFYLVQOTTHLMFMGYSMTAFCLPTMDKMLV
YSKIFLVGVFLSLFLIPYHKKMPVKERKRE"

BASE COUNT
ORIGIN
420 a 538 c 448 g 492 t

Query Match 54.3%; Score 1233.4; DB 10; Length 1698;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1404; Conservative 0; Mismatches 226; Indels 15; Gaps 2;

QY 48 GGGGGTCCCTGTGGGGCTCCCGAGTTAAGATGAGCTCTCCAGCGGAGGGGACGAGGG 107
| | | | |
Db 4 GGGGGTCCCGGGGGGGGGCCGAGTTAAGATGAGCTCTCCAGCGGAGGGGACGAGGG 63
| | | | |

QY 108 ACTGTGTGGCGCTGGCGGGGCTTGTGACGTGGGGTTTCCAGAGCTGAGCTTAAACA 167
| | | | |
Db 64 GAGACCTGTGAGCAGATGCGGGGGCTGTGGCGGGCTGTGAGATCTGAGCTTAAACA 123
| | | | |

QY 168 TTGGGAGCTCCCTGGGGGGCTGAGAACAGGGCTGGGGCTATCATCTCCATCTTCG 227
| | | | |
Db 124 TTGGGAGCTCTGGGGGGCTGAGAACAGGGCTGGGGCTATCATCTTCATCTTCG 183
| | | | |

QY 228 GGTACCCCTTGTGTTTATCCGACCTTCTTACAGGAGACCTTACCTTCATC 287
| | | | |
Db 184 GGTACCCCTTGTGTTTATCCGACCTTCTTACAGGAGACCTTACCTTCATC 243
| | | | |

QY 288 CACCTTTCATACCTTTACAGGCTCTCATATGCTTATTTTACCTTTGAAACGAGCTC 347
| | | | |
Db 244 CATCTTTCATACCTTTACAGGCTCTCATATGCTTATTTTACCTTTGAAACGAGCTC 303
| | | | |

QY 348 TACCACTCCCTGTGTATGTGCTTACAGTCTCATCTTCATCTTCATATGAGCGGAGC 407
| | | | |
Db 304 TACCACTCCCTGTGTATGTGCTTACAGTCTCATCTTCATCTTCATATGAGCGGAGC 363
| | | | |

QY 408 ATCACTGCGTCTCTCACTACCTTTTCTTCAGATGAGCTTACCTTGTGCTGATATAT 467
| | | | |
Db 364 GTCACTGCGTATTTACTACCTTTCTTCAGATGAGCTTACCTTGTGCTGATATAT 423
| | | | |

QY 468 TACACTGCAACGGCACTACATATCAAGTGAGCAATGACATGCTGTGCTGACTTTG 527
| | | | |
Db 424 TACACTGCAACGGCACTACATATCAAGTGAGCAATGACATGCTGTGCTGACTTTG 483
| | | | |

QY 528 AACCTATGTGCTTGTGCTGCTACTTGTGAGGAGGGAAGATACAGATTCCTGTGTC 587
| | | | |
Db 484 AACCTATGTGCTTGTGCTGCTACTTGTGAGGAGGGAAGATACAGATTCCTGTGTC 543
| | | | |

QY 588 TCTGAGCAAGAAATATGACATACGTGTGCTTCTTCCCTGTGAGAGTGTGCTGCTTC 647
| | | | |

Db 544 TCTGAGCAACAGAAATATGCCATACGGGGTTCCTTCATTGCTGGAAGTTGCTGCTTC 603
 Qy 648 TCTTACTTATAGGGGCTTCTTGTAGGGGCCCACTTCTCATATGATCATCATGAAG 707
 Db 604 TCTTACTTATAGGGGCTTCTTGTAGGGGCCCACTTCTCATATGATCATCATGAAG 663
 Qy 708 CTGGTTCAGGAGAGCTGATTTGACATACAGAAAGATACCAAGCATCATTCCTGCT 767
 Db 664 CTGGTTCAGGAGAGCTGATTTGACATACAGAAAGATACCAAGCATCATTCCTGCT 723
 Qy 768 CTCAAGCGCTGAGTGGGGCTTCTTCTACATAGTGGCTTACATCTGCTGACCCCTC 827
 Db 724 CTCAAGCGCTGAGTGGGGCTTCTTCTACATAGTGGCTTACATCTGCTGACCCCTC 783
 Qy 828 ATCACAGAAAGATATCTCTCTCACTGAGACTATGACCAACACCCCTTCTGTTCCGCTG 887
 Db 784 ATCACAGAGAGATATCTCTCTCACTGAGACTATGACCAACACCCCTTCTGTTCCGCTG 843
 Qy 888 ATGTACATCTGATCTGGGCAAGTTGTGTGTATACAAATATGTCACTGTGGCTGTC 947
 Db 844 ATGTACATCTGATCTGGGCAAGTTGTGTGTATACAAATATGTCACTGTGGCTGTC 903
 Qy 948 AAGAGAGAGATATGATTTTGAAGGGCTTGGCTTCAATGGCTTTGAAGAAAGGCAAG 1007
 Db 904 AAGAGAGAGATATGATTTTGAAGGGCTTGGCTTCAATGGCTTTGAAGAAATGGGACC 963
 Qy 1008 GCAAGTGGGATGCTGCTGCTCAACATGAGTGTGGCTCTTTGAACAAACCCCGCTTC 1067
 Db 964 GTGAGATGGGATGCTGCTGCTCAACATGAGTGTGGCTCTTTGAACAAACCCCGCTTC 1023
 Qy 1068 ACTGACCATTTGCTCTTCAATCAACACCAACGCTGGGTGGCCCGCTACATCTTC 1127
 Db 1024 AATGGACCATTTGCTCTTCAATCAACATCAATCAATGCTGGGTGGCCCGCTACATCTTC 1083
 Qy 1128 AAGAGATCAATTTCTTGAAGAAATGAAGTCTCTGAGGGTCTCTGTTCTTATCTTG 1187
 Db 1084 AAGAGATCAATTTCTTGAAGAAATGAAGTCTCTGAGGGTCTCTGTTCTTATCTTG 1143
 Qy 1188 GGCCTTGGGACGGCTGCTGCTGAGTATCTGCTGCTTCCAGATGGAATTCCTCAT 1247
 Db 1144 GGCCTTGGGACGGCTTACATCAGATCTGATTTGCTTCCAGATGGAATTCCTCAT 1203
 Qy 1248 GTTATTTGGAAGACAGGCTGCCAGCTCATTTCAAGAGAGCCCACTTGAGCAAGCTG 1307
 Db 1204 GTTATTTGGAAGACAGGCTGCCAGCTCATTTCAAGAGAGCCCACTTGAGCAAGCTG 1263
 Qy 1308 GCGGCTTACTGCTCTCAAGCCCTTCTTCTATTTGTTGCAACAGACATTCACCTGGCTC 1367
 Db 1264 GCGGCTTACTGCTCTCAAGCCCTTCTTCTATTTGTTGCAACAGACATTCACCTGGCTC 1323
 Qy 1368 TTCTATGGGTTACTCCATGATGCTTCTGCTTCTTCAAGTGGGACAAATGGCTTAAGTG 1427
 Db 1324 TTCTATGGGTTACTCCATGATGCTTCTGCTTCTTCAAGTGGGACAAATGGCTTAAGTG 1383
 Qy 1428 TATTAATTCATATTTCTTGGCCACATCTTCTTCTGAGCTTATATTTATTTGCT 1487
 Db 1384 TACAGATCATATTTCTTGGCAATGCTTCTTCTTGAAGCTTATATTTGCT 1443
 Qy 1488 TATATTCACAAAGCATGTGTGCCAAGAAAGAAAGTAAAGAGATGAATATC--- 1543
 Db 1444 TATATTCACAAAGCATGTGTGCCAAGAAAGAAAGTAAAGAGATGAATATC--- 1503
 Qy 1544 -CATTTCCCTGGTGGCTGTGGGAGCTGGTGCAGAAATACTGCTTCCCTTTTCAAG 1602
 Db 1504 TCTTTCCCTGGTGGCTGTGGGAGCTGGTGCAGAAATACTGCTTCCCTTTTCAAG 1563
 Qy 1603 CACTCTTTTGGCCCAAGAGCAAGATGGAAGAAAGCCAGGAGGTGGAAGATCATCTTC 1662
 Db 1564 CACTCTCTCAAGC-----AGCAGAGAGCGCAGAGAGTGAAGATCAGACGACTCC 1613
 Qy 1663 AGCTGTGCTTCTGCTGCAAGCAAG 1687
 Db 1614 AGCTGTGCTTCTGCTGCAAGCAAG 1638

RESULT 12
 AY028317
 LOCUS 1668 bp mRNA linear ROD 01-JUN-2001
 DEFINITION Mus musculus putative transmembrane protein PTG mRNA, complete cds.
 ACCESSION AY028317
 VERSION 1
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 1668)
 Zhu, Y., Han, Y. and Reddy, J. K.
 Cloning and initial characterization of mouse PTG cDNA, whose expression is in a PPAR alpha dependent manner
 Unpublished
 2 (bases 1 to 1668)
 Zhu, Y., Han, Y. and Reddy, J. K.
 Direct Submission
 Submitted (03-MAR-2001) Pathology, Northwestern University, 303 East Chicago Avenue, Chicago, IL 60611, USA
 JOURNAL
 TITLE
 AUTHORS
 REFERENCE
 1. 1968
 Location/Qualifiers
 FEATURES
 source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /chromosome="6"
 /tissue_type="liver"
 82..1545
 /note="PPAR alpha target"
 /codon_start=1
 /product="putative transmembrane protein PTG"
 /protein_id="AAK20915.1"
 /db_xref="GI:14278837"
 /translation="MASTADDMGSELTBOMRGLMPGVDELINLKATSLAGSEALRL
 IFSIFLGLPLAFYRHYI.PYKDSYI.IH.FHPTGISTAYPNFGHPSLSLCLVQL
 IRLMGRVAVITTLCPQMAVLLAGVYVATGDDIDKMTWPHCVLTLLKLGCTIDY
 DGGKGNLSLTSBQKSLAIVGSLVLAAGVLSFVYAFVLPQFSNMHWKLVKGLTD
 IPGKPNSTIPLKLSLGLVLYVLTSLPHITDYLTDYDNR.PMPFRMGLIIV
 GKFVLYKVTGCMVLTGEGVILSGFNGFNGEIVRMQACAMKMYL.PETTPRNGTI
 ASPNNTAMVARYI.FRLKPLGNELSGSLFLALMHLGHSYLICPQWELIIV
 VEKQVSLIRSPALSLASITLADLPFIYLVQOTTHWLPNGYSMTAFCLFTWDMKLK
 YRSIFYLGHVFFLSLFTLPIYIHKMVRKELKRE"
 BASE COUNT 443 a 554 c 469 g 502 t
 ORIGIN
 Query Match 54.1%; Score 1229.4; DB 10; Length 1668;
 Best Local Similarity 84.6%; Pred. No. 0;
 Matches 1433; Conservative 0; Mismatches 241; Indels 19; Gaps 4;
 Qy 1 GGGGGTGAAGCATTCGTTTGGCCCGCATTCGGGGCGCGCGGATGCGGGGGTCCCTG 60
 Db 8 GGGGTGAAGCACCCTGCTTCCCAATTCGGGGCGCGGAG---CTGGGGGTCCCCGG 64
 Qy 61 GGGCTCCGGAGTTAAGATGCGCTCTCAGCGGAGGGGAGCAGAGGAGCTGTGTGGCG 120
 Db 65 GGGCGCCCGCAGTTAAGATGCGCTCTCAGCGGAGGGGAGCAGTGGAGAGCGTGGAGC 124
 Qy 121 TGGCGGGGCTTGTGAGTGGGTTTCCAGAGCTAGAGCTTAAAGATTTGGAGAGTCCC 180
 Db 125 AGATGCGGGGCTGTGGCGGGGTGTGAGGATCTGAGCTTAAAGATTTGGAGAGTCTC 184
 Qy 181 TGGGCGCTCAGAACAGCGCTGCGGCTGATCATCTCTCTGGGTTACCCCTTTG 240
 Db 185 TGGGCGCTCAGAACAGCGCTGCGGCTGATCATCTCTCTGGGTTACCCCTTTG 244
 Qy 241 CTTTGTATTATGCGATTACTTTTCAAGAGAGACCTTACCTATCCACCTTTTCATA 300
 Db 245 CTCTGTTTACCGGATTAACCTTTTCAAGAGAGACCTTACCTATCCATCTCTTCACA 304

QY 301 CCTTACAGGCTTCATGCTTATTTTAACTTGGAAACAGCTCTACCACTCCCTCG 360
 DB 305 CCTTACAGGCTTCATGCTTATTTTAACTTGGAAACAGCTCTACCACTCCCTCG 364
 QY 361 TGTGATGCTTACGCTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 420
 DB 365 TATGTGCGTCTTCACTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 424
 QY 421 TCACTACCTTCTGCTTCAAGATGCTTCTGCTGATCTATTTACATGCGCAAG 480
 DB 425 TTAATACCTTCTGCTTCAAGATGCTTCTGCTGATCTATTTACATGCGCAAG 484
 QY 481 GCAATACGATTCATGATGCAATGCAATGCTTCTGCTGATCTATTTACATGCGCAAG 540
 DB 485 GTGATACGATTCATGATGCAATGCAATGCTTCTGCTGATCTATTTACATGCGCAAG 544
 QY 541 TGTGATGCTTACGCTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 600
 DB 545 TGTGATGCTTACGCTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 604
 QY 601 AATATGCAATGCTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 660
 DB 605 AATATGCAATGCTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 664
 QY 661 GGGCTTCTGATGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 720
 DB 665 GAGCTTCTTGTGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 724
 QY 721 AGCTGATGCAATGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 780
 DB 725 AGCTGATGCAATGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 784
 QY 781 GTTGGGCTTCTGATGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 840
 DB 785 GTTGGGCTTCTGATGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 844
 QY 841 ATCTGCTCACTGATGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 900
 DB 845 ATCTGCTCACTGATGAGGCGCCAGTCTCAATGATCACTACATGAGCTGTCAGGAG 904
 QY 901 TCTGGGCAAGTGTGCTTCAATGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 905 TCTGGGCAAGTGTGCTTCAATGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
 QY 961 GCAATTTGAGGCGCTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1020
 DB 965 GCAATTTGAGGCGCTTCTCATCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1024
 QY 1021 CCTGTGCAATGAGGCGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1080
 DB 1025 CCTGTGCAATGAGGCGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1084
 QY 1081 CCTGATTTCAATGAGGCGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1140
 DB 1085 CCTGATTTCAATGAGGCGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1144
 QY 1141 TCTTGGAAATTAAGAACTCTTCAAGGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1200
 DB 1145 TCTTGGAAATTAAGAACTCTTCAAGGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1204
 QY 1201 GCTGTGCAATGAGGCGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1260
 DB 1205 GCTGTGCAATGAGGCGCTTCTGCTTAAATGAGCCGACCATCTGCGCTC 1264
 QY 1261 GACAGGCTGCAAGGCTTCTCAAGAGGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCT 1320
 DB 1265 AGAGAGTCAAGGCTTCTCAAGAGGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCT 1324
 QY 1321 TCTTCCAGGCTTCTCAATGAGGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 DB 1325 TCTTCCAGGCTTCTCAATGAGGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1384
 QY 1381 CCAATGAGGCTTCTGCTTCAAGGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1440

DB 1385 CTATGACTCTTCTGCTTCTGCTTCAATGAGGCGCCAGTCTGCTGCTGCTGCTGCT 1444
 QY 1441 ATTTCTTGGCAATGCTTCTTCTGCTTCAATGAGGCGCCAGTCTGCTGCTGCTGCT 1500
 DB 1445 ATTTCTTGGCAATGCTTCTTCTGCTTCAATGAGGCGCCAGTCTGCTGCTGCTGCT 1504
 QY 1501 CAATGCTGCAAGGAAAGGTTTAAAGATGGAATTC-----CATTTCCCTGCT 1555
 DB 1505 CAATGCTGCAAGGAAAGGTTTAAAGATGGAATTC-----CATTTCCCTGCT 1564
 QY 1556 GGGCTTGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1614
 DB 1565 GCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1624
 QY 1615 CCAGAGCAGAGATGAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1674
 DB 1625 C-----AGCAGAGAGGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1674
 QY 1675 GCTGCGCAGCAG 1687
 DB 1675 GCGAGAGCAGCAG 1687

RESULT 13
 BT007000
 LOCUS
 DEFINITION
 Homo sapiens putative protein similar to nesey (Drosophila) mRNA,
 complete cds.
 ACCESSION
 BT007000
 VERSION
 BT007000.1 GI:30582838
 KEYWORDS
 FLI_CDN,
 Homo sapiens (human)
 SOURCE
 Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1146)
 Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
 Pheilan,M., and Farmer,A.
 TITLE
 Cloning of human full-length CDSs in BD Creator(TM) System Donor
 vector
 JOURNAL
 Unpublished
 2 (bases 1 to 1146)
 Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
 Pheilan,M., and Farmer,A.
 AUTHORS
 Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
 Pheilan,M., and Farmer,A.
 COMMENT
 Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
 Circle, Palo Alto, CA 94303, USA
 This CDS clone is a part of a collection of human full length
 expression clones generated by BD Biosciences Clontech and the
 Harvard Institute of Proteomics. Each CDS has been cloned in two
 forms: with and without stop-codon (to allow fusion with C-terminal
 tag). The CDS has been directionally cloned using BD In-Fusion(TM)
 cloning system between the SalI and HindIII sites of the pDNR-DUAL
 vector. Additional sequences in the clone: 'ACC' after SalI site
 and before 'ATG' to provide Kozak consensus sequence; 'GG' after
 last codon and before HindIII site to maintain reading frame.
 Clone distribution: <http://bioinfo.clontech.com/orfclones>.
 Location/Qualifiers
 1..1146
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="GH00356X1.0"
 /clone_1kb="BD Creator(TM) CDS Library derived from MGC
 collection"
 /lab_host="DH5alpha T1 resistant"
 /note="Vector: pDNR-Dual"
 1..1146
 /product="putative protein similar to nesey (Drosophila)"

```

/Protein_id="AAP35646.1"
/db_xref="GI:30582839"
/translation="MGRITRAVLTFPCFQMAVLLAGYYTATNDYIKWMPHCVLLT
KILGLAVDYEDGKDONSLSSBOOKYXGIVLEAVPSYVYGAFLVGPDSNMHY
MLVQVELIDIDPKIENSIIPAKRISLGLFLVGTLLSPHITERYLLTEDNDHF
WRCHMMLMKRFLVLYKYTCMLVTGVCILTELGNGBEKKAAWDCANNKWL
ETNPRFTGTLASNTNANVARYIFKRFLGNKELSGSLFLMALHNGHSGYLV
CFQMKFLIVIERQOARLJOSEFTLSKLAITVLQPFYLVQOTLHLMFGHYSMTAF
LFTWDMKLVKYSIYFLGHIFFLSLFLPIYIKHAWVPRKRLKKE"
BASE COUNT      274 a      310 c      256 g      306 t
ORIGIN
Query Match      50.3%; Score 1143.4; DB 9; Length 1146;
Best Local Similarity 99.9%; Pred. No. 1.7e-299;
Matches 1144; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 396 ATGGGGCCGACCATCACTGCGGCTCTCACTACTTTTGGTTCAGATGGCTACCTTCTG 455
DB 1 ATGGGGCCGACCATCACTGCGGCTCTCACTACTTTTGGTTCAGATGGCTACCTTCTG 60
QY 456 GCTGATACCTATTACACTGCGACCGGCACTACATATCAAGTGCACATGCCACATGTT 515
DB 61 GCTGATACCTATTACACTGCGACCGGCACTACATATCAAGTGCACATGCCACATGTT 120
QY 516 GTTCTGACTTTGAAGCTGATGTTGGCTGTTGACTTCTTGAACGAGGAGAAAGATCAG 575
DB 121 GTTCTGACTTTGAAGCTGATGTTGGCTGTTGACTTCTTGAACGAGGAGAAAGATCAG 180
QY 576 AATTCTCTTCTCTTGAAGCAAGAAATATGCATACGTTGTTCTTCTTCTCTGCTGAA 635
DB 181 AATTCTCTTCTCTTGAAGCAAGAAATATGCATACGTTGTTCTTCTTCTCTGCTGAA 240
QY 636 GTTGTGCTGTTCTCTACTTCTTATGAGGCGCTTCTTGTAGGCGCCAGTTTCAATGAT 695
DB 241 GTTGTGCTGTTCTCTACTTCTTATGAGGCGCTTCTTGTAGGCGCCAGTTTCAATGAT 300
QY 696 CACTACATGAAGCTGTGACGAGGAGAGCTGATGACATACAGAAAGATACCAACACG 755
DB 301 CACTACATGAAGCTGTGACGAGGAGAGCTGATGACATACAGAAAGATACCAACACG 360
QY 756 ATCATTCCTGCTCTCAAGCGGCTGAGTCTGGGCTTTTCTACTAGTGGCTTACACATG 815
DB 361 ATCATTCCTGCTCTCAAGCGGCTGAGTCTGGGCTTTTCTACTAGTGGCTTACACATG 420
QY 816 CTCAGCGCCCAATCAAGAAAGCTATCTCTCACTGAAAGATATACCAACACCGCCCTTC 875
DB 421 CTCAGCGCCCAATCAAGAAAGCTATCTCTCACTGAAAGATATACCAACACCGCCCTTC 480
QY 876 TGGTTCGCTGATGATGCTGATCTGGGCAAGTTTGTGCTGTACAATAATGTCAAC 935
DB 481 TGGTTCGCTGATGATGCTGATCTGGGCAAGTTTGTGCTGTACAATAATGTCAAC 540
QY 936 TGTGGCTGTGTACAGAAAGATGATGATTTGACCGGCTTGGGCTTCAATGGCTTTGAA 995
DB 541 TGTGGCTGTGTACAGAAAGATGATGATTTGACCGGCTTGGGCTTCAATGGCTTTGAA 600
QY 996 GAAAAGGAGGAGGAGGAGTGTGCTGTGGCAACATGAAGTGTGGCTTTTGAAGA 1055
DB 601 GAAAAGGAGGAGGAGGAGTGTGCTGTGGCAACATGAAGTGTGGCTTTTGAAGA 660
QY 1056 AACCCCGCTTCACTGACATGCTGCTATTCATCAATCAACACCAACCGCTGGTGGCC 1115
DB 661 AACCCCGCTTCACTGACATGCTGCTATTCATCAATCAACACCAACCGCTGGTGGCC 720
QY 1116 CGCTCATTTTCAAAAGACTCAAGTTCTTTGAAAATAAGAACTCTCTCAGGGTCTCTG 1175
DB 721 CGCTCATTTTCAAAAGACTCAAGTTCTTTGAAAATAAGAACTCTCTCAGGGTCTCTG 780
QY 1176 TTGCTATTTCTGGGCTTGGACCGGCTGACACTAGATATACGTTGCTGCTTCCAGATG 1235
DB 781 TTGCTATTTCTGGGCTTGGACCGGCTGACACTAGATATACGTTGCTGCTTCCAGATG 840
QY 1236 GAATTCCTATTGTTATTTGTGAAGAAGAGGCTGCCAGGCTCATTCAGAGAGCCCCAC 1295

```

```

DB 841 AATTCCTATTGTTATTTGTGAAGAAGAGCGGCTGCCAGGCTCATTTAAGAGAGCCCCAC 900
QY 1296 CTGAGAACCTGCGCCCATTAATCTGCTCTCCAGGCTTTACTATTGTTGTCACAGACC 1355
DB 901 CTGAGAACCTGCGCCCATTAATCTGCTCTCCAGGCTTTACTATTGTTGTCACAGACC 960
QY 1356 ATCCACTGCTCTTCAATGAGTTACTCCAGATGCGCTTGGCTCTTCAAGTGGGACAA 1415
DB 961 ATCCACTGCTCTTCAATGAGTTACTCCAGATGCGCTTGGCTCTTCAAGTGGGACAA 1020
QY 1416 TGGCTTAAGTGTATTAATCCATCTATTCTTGGCCACATCTTCTCTGAGCCCTACTA 1475
DB 1021 TGGCTTAAGTGTATTAATCCATCTATTCTTGGCCACATCTTCTCTGAGCCCTACTA 1080
QY 1476 TTCAATTTGCTTATTTATTCACAAAGCAATGTGTCCAGAGAAAGAGTTAAAGATG 1535
DB 1081 TTCAATTTGCTTATTTATTCACAAAGCAATGTGTCCAGAGAAAGAGTTAAAGATG 1140
QY 1536 GAATA 1540
DB 1141 GAATA 1145
RESULT 14
BT007735 1146 bp mRNA linear SYN 13-MAY-2003
LOCUS BT007735
DEFINITION Synthetic construct Homo sapiens putative protein similar to nesey
(Drosophila) mRNA, partial cds.
ACCESSION BT007735
VERSION BT007735.1 GI:30584308
KEYWORDS FLI CDNA.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1146)
AUTHORS Kauline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Pheilan,M. and Farmer,A.
Cloning of human full-length CDSs in BD Creator(TM) System Donor
vector
TITLE Unpublished
JOURNAL 2 (bases 1 to 1146)
REFERENCE Kauline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Pheilan,M. and Farmer,A.
Direct Submission
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA
COMMENT This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD in-Fusion(TM)
cloning system between the SalI and HindIII sites of the pMDR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ATG' to provide Kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.
Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES
source
1..1146
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="G000356L1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/notice="vector: pMDR-Dual"
1..>1146
/notice="Mutations: 1145:Stop->Leu"
/codon_start=1
/transl_table=11

```

/product="Homo sapiens putative protein similar to nesy
(Drosophila)"
/protein_id="AA936403.1"
/db_xref="GI:30584309"
/translation="MGRTTAVLTTFPCFQAVLLAAGYYTATGNVDIKMTMPCVLT
KLGLAVDYGDKDQSLSSBOQKXARVPSLLLEAGASYGAFVLPQSCMHY
MKLVQELIDIPKIPNSIIIPALRSLISFYLVGYTLSPHLEDTLEDNDHF
WFRQWMLIWGKFLVYKVTGCVLTGSGVCLTGLGPNPEKPKARMDACAMKWL
ENMPRTGTASPNININAMVARYIPRKLFLNKEISQSLPLALNHLGSLV
CFQMKFLIVERQARLQESFPLSLAIVLPPIYVQDTIHMLFGYSMTAF
LFTWDMKLVKYSIYFLGHIFFLSLFIPIYHKAVPRKXKMKEL"

BASE COUNT 273 a 310 c 256 g 307 t

ORIGIN

Query Match 50.3%; Score 1142.4; DB 12; Length 1146;
Best Local Similarity 99.9%; Pred. No. 3.2e-299;
Matches 1143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 396 ATGGGCGGACACATCACTGCGGCTCTCACTACCTTTGGTCCAGATGGCTACCTCTG 455
DB 1 ATGGGCGGACACATCACTGCGGCTCTCACTACCTTTGGTCCAGATGGCTACCTCTG 60
QY 456 GCTGATATCTATTACATCTGCCACCGGCACTAGCATATCAAGTGAACATGCCATTTG 515
DB 61 GCTGATATCTATTACATCTGCCACCGGCACTAGCATATCAAGTGAACATGCCATTTG 120
QY 516 GTTCTGACTTTGAAGCTGATTTGGTGGCTGTGACTACTTTTGAAGGAGGAAAGATCAG 575
DB 121 GTTCTGACTTTGAAGCTGATTTGGTGGCTGTGACTACTTTTGAAGGAGGAAAGATCAG 180
QY 576 AATTCCCTGTCTCTGAGCAACAGAAATATGCCATAGTGTGTCTCTCTGCTGAGAA 635
DB 181 AATTCCCTGTCTCTGAGCAACAGAAATATGCCATAGTGTGTCTCTCTGCTGAGAA 240
QY 636 GTTGTGTGTCT 695
DB 241 GTTGTGTGTCT 300
QY 696 CACTTACATGAAGCTGTGAGGAGGAGAGCTGATGACATACAGAGAAAGATACCAACAGC 755
DB 301 CACTTACATGAAGCTGTGAGGAGGAGAGCTGATGACATACAGAGAAAGATACCAACAGC 360
QY 756 ATCATTCCTCTCTCAAGCCCTGAGTGTGGGCTTTTCTACCTAGTGGGCTACACACTG 815
DB 361 ATCATTCCTCTCTCAAGCCCTGAGTGTGGGCTTTTCTACCTAGTGGGCTACACACTG 420
QY 816 CTCAGCCCCACATCACAGAAAGCTATCTCTCACTGAGACTATGACCAACCCCTTC 875
DB 421 CTCAGCCCCACATCACAGAAAGCTATCTCTCACTGAGACTATGACCAACCCCTTC 480
QY 876 TGGTTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 935
DB 481 TGGTTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 936 TGTGTGCTGTCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 995
DB 541 TGTGTGCTGTCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 996 GAAAGGCGCAAGGCAAGTGGATGCTGTGCCAATGAGAGTGTGGCTCTTTGAACA 1055
DB 601 GAAAGGCGCAAGGCAAGTGGATGCTGTGCCAATGAGAGTGTGGCTCTTTGAACA 660
QY 1056 AACCCCGGCTTCACTGGACATTTGCTCATTTCAACATCAACACCAAGCCCTGGGTGGCC 1115
DB 661 AACCCCGGCTTCACTGGACATTTGCTCATTTCAACATCAACACCAAGCCCTGGGTGGCC 720
QY 1116 CGGTACATCTTCAAGACACTCAAGTCTCTTGAAGAAATGATGATGATGATGATGATG 1175
DB 721 CGGTACATCTTCAAGACACTCAAGTCTCTTGAAGAAATGATGATGATGATGATGATG 780
QY 1176 TTGCTATTCTGGCCCTCTGAGCAAGGCTGACCTACAGAGATACCTGTCTGCTTCAAG 1235
DB 781 TTGCTATTCTGGCCCTCTGAGCAAGGCTGACCTACAGAGATACCTGTCTGCTTCAAG 840

QY 1236 GAATTCCTCATTTCTTATTGTGGAAGACAGAGCTGCCAGGCTCATTTCAAGAGACCCACC 1295
DB 841 AAATTCCTCATTTCTTATTGTGGAAGACAGAGCTGCCAGGCTCATTTCAAGAGACCCACC 900
QY 1296 CTGAGCAAGCTGGCCGCCCATTTACTGTCTCCAGCCCTTCTACTATTGTGCAACAGACC 1355
DB 901 CTGAGCAAGCTGGCCGCCCATTTACTGTCTCCAGCCCTTCTACTATTGTGCAACAGACC 960
QY 1356 ATCCACTGCTCTTCAATGGGTTACTCCATGATGCTGCTTCTGCTTCAAGTGGACAAA 1415
DB 961 ATCCACTGCTCTTCAATGGGTTACTCCATGATGCTGCTTCTGCTTCAAGTGGACAAA 1020
QY 1416 TGGCTTAAGGTATTAATCAATCTTCTCTTGGCCACATCTTCTCTGAGCCCTACAT 1475
DB 1021 TGGCTTAAGGTATTAATCAATCTTCTCTTGGCCACATCTTCTCTGAGCCCTACAT 1080
QY 1476 TTGATATTCCTTATATTCACAAAGCATGTGTCACAAAGAGAGATTAAGAAGATG 1535
DB 1081 TTGATATTCCTTATATTCACAAAGCATGTGTCACAAAGAGAGATTAAGAAGATG 1140
QY 1536 GAAT 1539
DB 1141 GAAT 1144

RESULT 15
AC006512/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 155975) bp DNA linear PRI 04-APR-2003
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barberia,J., Benton,J., Bimaye,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Burch,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chin,D., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
Gill,R., Gorell,J.H., Guevara,M., Gunaratne,P., Hale,S.,
Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Hawla,P.,
Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M.,
Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulik,S.,
Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
King,L., Korvah,J., Kovar,C., Kratochvil,J., Kureishi,A., Landry,N.,
Leal,B., Lee,E., Lewis,L.C., Lewis,L.J., Li,J., Li,Z., Licharge,O.,
Lieu,C., Liu,J., Liu,W., Loulsegé,H., Lozano,R.J., Lu,X.,
Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
Marandel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Merscher,S.,
Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pui,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shim,C., Shooshitari,N., Sisson,I.,

Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telirod, B., Thomas, N., Thomas, S.,
 Usmani, K., Vaequez, L., Vera, V., Valladon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Kuchelapatti, R.,
 Weinstein, G. and Gibbs, R.

TITLE
 Direct Submission

JOURNAL
 Unpublished

REFERENCE
 2 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (05-FEB-1999) Molecular and Human Genetics, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 3 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (29-MAY-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 4 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (02-JUN-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 5 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (03-JUN-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 6 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (18-JUN-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 7 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (27-OCT-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 8 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (29-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 9 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (02-APR-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 10 (bases 1 to 155975)

AUTHORS
 Worley, K.C.

TITLE
 Direct Submission

JOURNAL
 Submitted (04-APR-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT
 On Apr 2, 2003 this sequence version replaced gi:4926863.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:
 STS are identified using ePCR (genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.
 Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished) for Human and Mouse sequences.
 Genes and Region of sequence similarity are identified by BLAST
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
 standards - estimated error rate less than 1 per 10,000 bases.
 Reports of lowest quality individual bases and measures of base
 quality are listed below. Description of the metrics can be found
 at URL:
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

source

Location/Qualifiers

1..155975
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="12"
 /clone="RP3-461P17"

misc_feature

1..97210
 /note="overlaps bases 127585..222930 of clone"

repeat_region

1649..1772
 /function="clone overlap"

repeat_region

1772..1772
 /rpt_family="13b"

repeat_region

complement(2781..3065)
 /rpt_family="AluYd2"

repeat_region

3127..3430
 /rpt_family="AluX"

STS

3473..4048
 /standard_name="GDB:214837"

repeat_region

3592..3797
 /rpt_family="MIR"

repeat_region

4550..4771
 /rpt_family="13b"

repeat_region

complement(4816..5127)
 /rpt_family="AluSp"

repeat_region

complement(5146..5288)
 /rpt_family="FLM_C"

repeat_region

complement(5322..5506)
 /rpt_family="AluX"

repeat_region

5541..5833
 /rpt_family="Aluub"

STS

6770..6912
 /standard_name="GDB:384728"

STS

7448..7601
 /standard_name="D12S1899"

STS

7675..7889
 /standard_name="G10506"

repeat_region

7717..7746
 /rpt_family="(TTTA)n"

STS

7769..7860
 /standard_name="SHGC-58828"

repeat_region

9529..9570
 /rpt_family="(TG)n"

repeat_region

complement(10903..11196)
 /rpt_family="AluX"

repeat_region

12008..12158


```
repeat_region      /rpt_family="CT_rich"  
                    1216f, 12206  
                    /rpt_family="rTCGCC" n"  
repeat_region      1236f, 12555  
                    /rpt_family=" (CGGG) n"  
repeat_region      12647, 12655  
                    /rpt_family="GC_rich"  
repeat_region      12822, 12830  
                    /rpt_family="GC_rich"
```

Query Match	31.2%	Score 708.6;	DB 9;	Length 155975;
Best Local Similarity	99.4%;	Pred. No. 8.7e-181;		
Matches 711;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

Qy	1546	TTTCCCTGGAGGCGCTGTGCGGGAAGTGGTGAGAAAATACTCGCTCCCTTTTTCACAGCAC	1605
Db	61329	TCTTTCAGAGTGGCTGTGCGGGAAGTGGTGAGAAAATACTCGCTCCCTTTTTCACAGCAC	612707
Qy	1606	TCCTTTGCCCCAGAGCAGAGATGGA AAAACCCAGGGAGGTGGAAGATCGATGCTTCACG	1665
Db	61269	TCCCTTTGCCCCAGAGCAGAGATGGA AAAACCCAGGGAGGTGGAAGATCGATGCTTCACG	612110
Qy	1666	TGTCCTCTGTGTGCGAGCCAGATCTTCAATTGGGGGCCAAAAGGGGAAACTTTTTTTGGAG	1725
Db	61209	TGTCCTCTGTGTGCGAGCCAGATCTTCAATTGGGGGCCAAAAGGGGAAACTTTTTTTGGAG	611500
Qy	1726	AAGCGTCTTGCTTTGTGATCCCAAGCTGGAATGATGCGGGATCTCAAGTTCACCGCAA	1785
Db	61149	AAGCGTCTTGCTTTGTGATCCCAAGCTGGAATGATGCGGGATCTCAAGTTCACCGCAA	610950
Qy	1786	CCTCACCCTCTGGGGTTCAAGTATATTTTCTGGCTCAAGCCTCCCAAGTATGCTGGGAATAC	1845
Db	61089	CCTCACCCTCTGGGGTTCAAGTATATTTTCTGGCTCAAGCCTCCCAAGTATGCTGGGAATAC	610300
Qy	1846	AGGCAGCCCAACATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGATTTTCACACAG	1905
Db	61029	AGGCAGCCCAACATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGATTTTCACACAG	609707
Qy	1906	TTGGCCAGGCTGGTCTTGAACTCTTACCGCAATGATCCAACCGGCTCGGCTCCCAAA	1965
Db	60969	TTGGCCAGGCTGGTCTTGAACTCTTACCGCAATGATCCAACCGGCTCGGCTCCCAAA	60910
Qy	1966	GTGCTGGGAATTCACAGCGTAGACCAGCTGCCGGCCCAAAAGGGGAAACTCTGTGGAG	2025
Db	60909	GTGCTGGGAATTCACAGCGTAGACCAGCTGCCGGCCCAAAAGGGGAAACTCTGTGGAG	60850
Qy	2026	GAGCAGAGGGGCTACACTTCCTCCCTCTGATTTCCGCCATGCACTTGCCTTATCTCTCCCC	2085
Db	60849	GAGCAGAGGGGCTACACTTCCTCCCTCTGATTTCCGCCATGCACTTGCCTTATCTCTCCCC	607900
Qy	2086	ATTACGACAGGAATCTATTTGTTTCTTCTGCGCAATTTACTATGATTTGTGATGTCG	2145
Db	60789	ATTACGACAGGAATCTATTTGTTTCTTCTGCGCAATTTACTATGATTTGTGATGTCG	60730
Qy	2146	GCTAACCAACACCCCCCATGAGGGGGGTGAGAGAGGGGTGCAAGGCTCTGCTCCACT	2205
Db	60729	GCTAACCAACACCCCCCATGAGGGGGGTGAGAGAGGGGTGCAAGGCTCTGCTCCACT	60670
Qy	2206	TTTTCTACCTTGGAACTGTATTAATTAATAATCACTCTGTGTTGTCAAGTTTTC	2260
Db	60669	TTTTCTACCTTGGAACTGTATTAATTAATAATCACTCTGTGTTGTCAAGTTTTC	60615

Search completed: November 22, 2003, 05:56:24
Job time : 8298 secs

PR 14-MAY-1999; 99US-0311937.
PR 14-MAY-1999; 99US-0311940.
XX
PA (INCY-) INCYTE GENOMICS INC.
PI Yue H, Tang YF, Lal P, Reddy R, Batria S, Baughn MR, Yang J;
PI Azimtai Y, Lu DM, Au-Young J, Shih LL;
DR MPI; 2001-016234/02.
XX P-PSDB; AAB36584.
PT Human FLEXHT protein and DNA sequences, useful for treating
PT immunological disorders, developmental disorders, and cancers -
XX
XX Claim 5; Page 140; 168pp; English.
XX AAC8070 to AAC8124 encode the 55 FLEXHT (full-length molecules
CC expressed in human tissues) proteins given in AAB36579 to AAB36633. The
CC present invention describes an isolated polypeptide (A) comprising an
CC amino acid sequence selected from one of 55 amino acid sequences 42-876
CC residues in length, corresponding to FLEXHT-1 to FLEXHT-55, a 90 %
CC identical sequence, and a biologically active or immunogenic fragment of
CC the sequence. The FLEXHT proteins can have antianaemic, anticonvulsant,
CC antiarteriosclerotic, immunomodulatory, cytostatic, antiasthmatic,
CC antiinflammatory, hepatotropic, antidiabetic, anti-gout, antihypertoid,
CC neuroprotective, antiallergic, osteopathic, antipsoriatic, antitumor
CC and antirheumatic activities, and can be used in gene therapy. The
CC polynucleotide sequences can be used to express the protein sequences.
CC Pharmaceutical compositions comprising FLEXHT can be used to treat
CC diseases or conditions associated with altered expression of functional
CC FLEXHT. The proteins and polynucleotides can be used to diagnose and
CC treat disorders including anaemia, epilepsy, arteriosclerosis,
CC atherosclerosis, developmental disorders, cancers, and immunological
CC disorders such as asthma, bronchitis, cirrhosis, Crohn's disease,
CC diabetes mellitus, gout, Grave's disease, multiple sclerosis,
CC osteoarthritis, pancreatitis, psoriasis, rheumatoid arthritis, and
CC ulcerative colitis.
XX
XX Sequence 2273 BP; 494 A; 630 C; 547 G; 602 T; 0 other;
SQ
Query Match 100.0%; Score 2273; DB 22; Length 2273;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGGTGAAGCGATACGTTTGGCCCGCATTCGCGCGCGGAGCTGGGGGGTCCCTGG 60
DB 1 GGGGGTGAAGCGATACGTTTGGCCCGCATTCGCGCGCGGAGCTGGGGGGTCCCTGG 60
QY 61 GGGCTCCCGAGTTAAGATGGCGTCTCAAGCGAGGGGAGCGAGGGAGCTGTGTGGCG 120
DB 61 GGGCTCCCGAGTTAAGATGGCGTCTCAAGCGAGGGGAGCGAGGGAGCTGTGTGGCG 120
QY 121 TGGCGGGGGTTCGACGTGGGTTTCCAGAGAGCTGAACCTTAACAAGTTGGCAGCTCC 180
DB 121 TGGCGGGGGTTCGACGTGGGTTTCCAGAGAGCTGAACCTTAACAAGTTGGCAGCTCC 180
QY 121 TGGCGGGGGTTCGACGTGGGTTTCCAGAGAGCTGAACCTTAACAAGTTGGCAGCTCC 180
DB 121 TGGCGGGGGTTCGACGTGGGTTTCCAGAGAGCTGAACCTTAACAAGTTGGCAGCTCC 180
QY 181 TGGCGCGGTGAGAAAGCGGCTGGCGTGAATCTTCATCTTCTGGGTTACCCCTTTG 240
DB 181 TGGCGCGGTGAGAAAGCGGCTGGCGTGAATCTTCATCTTCTGGGTTACCCCTTTG 240
QY 241 CTCTGTTTATGGGCACTTACCTTTCTCAAGAGAGCTACCTCATCTTCACCTTCCTCA 300
DB 241 CTCTGTTTATGGGCACTTACCTTTCTCAAGAGAGCTACCTCATCTTCACCTTCCTCA 300
QY 301 CCTTTACAGGCTCTCAATGCTTAATTTAATTTGAAACAGCTCTACACCTCCCTGC 360
DB 301 CCTTTACAGGCTCTCAATGCTTAATTTAATTTGAAACAGCTCTACACCTCCCTGC 360
QY 361 TGTGTATGTGCTTCACTTCTCATCTTCTGACTAATGGGCGGACCATCATGCGCTCC 420
DB 361 TGTGTATGTGCTTCACTTCTCATCTTCTGACTAATGGGCGGACCATCATGCGCTCC 420
QY 421 TCACCTACCTTTGCTTCCGAATGGCCCTTCTGGTGATACTATTACATGCGCCG 480

DB 421 TCACCTACCTTTGCTTCCGAATGGCCCTTCTGGTGATACTATTACATGCGCCG 480
QY 481 GCACTACGATATCAAGTGGACATGCCATTTGTTCTGACTTTGAAGCTGATTTGTT 540
DB 481 GCACTACGATATCAAGTGGACATGCCATTTGTTCTGACTTTGAAGCTGATTTGTT 540
QY 541 TGGCTGTGACTTACTTTGACGAGGAGAAAGATCAGATTTCTTGTCTTGAGCAACA 600
DB 541 TGGCTGTGACTTACTTTGACGAGGAGAAAGATCAGATTTCTTGTCTTGAGCAACA 600
QY 601 AATATGACCTAGTGTGTTCTTCCCTGCTGAGAGTGTGCTGCTTCTCTCTTATG 660
DB 601 AATATGACCTAGTGTGTTCTTCCCTGCTGAGAGTGTGCTGCTTCTCTCTTATG 660
QY 661 GGGCTCTTGTGAGGGGCCCATGTTCTCAATGATCTCAATGAGAGCTGTGACAGGAG 720
DB 661 GGGCTCTTGTGAGGGGCCCATGTTCTCAATGATCTCAATGAGAGCTGTGACAGGAG 720
QY 721 AGCTGATGACATACCAAGAAAGATACCAACAGCATCTTCTGCTTCAAGCGCTGA 780
DB 721 AGCTGATGACATACCAAGAAAGATACCAACAGCATCTTCTGCTTCAAGCGCTGA 780
QY 781 GTCTGGGCTTTTCTACCTAGTGGGTACACCTGCTGAGCCGCCCATCACAGAGACT 840
DB 781 GTCTGGGCTTTTCTACCTAGTGGGTACACCTGCTGAGCCGCCCATCACAGAGACT 840
QY 841 ATCTCTCTCACTAGAGATGACCAACCAACCCCTTGTGTTCCGCTGATGATGCTGA 900
DB 841 ATCTCTCTCACTAGAGATGACCAACCAACCCCTTGTGTTCCGCTGATGATGCTGA 900
QY 901 TCTGGGGCAAGTTTGTGCTGTACAAATATGTCACCTGTGCTGCTGACAGAGAGTAT 960
DB 901 TCTGGGGCAAGTTTGTGCTGTACAAATATGTCACCTGTGCTGCTGACAGAGAGTAT 960
QY 961 GCATTTTGAAGGGGCTGCTCAATGCTTGAAGAAAGGAGCAAGGAGTGGAGT 1020
DB 961 GCATTTTGAAGGGGCTGCTCAATGCTTGAAGAAAGGAGCAAGGAGTGGAGT 1020
QY 1021 CCTGTGCAACATGAGAGTGTGGCTCTTTGAAGCAAAACCCCGCTTCACTGGACCATTTG 1080
DB 1021 CCTGTGCAACATGAGAGTGTGGCTCTTTGAAGCAAAACCCCGCTTCACTGGACCATTTG 1080
QY 1081 CCTCATTTCAACATCAACCAACAGCTGGGTGGCCGCTACATCTTCAAAAGCATCAAGT 1140
DB 1081 CCTCATTTCAACATCAACCAACAGCTGGGTGGCCGCTACATCTTCAAAAGCATCAAGT 1140
QY 1141 TCTTGGAAATTAAGAACTCTCAAGAGTCTCTGCTGCTATTTCTGGCCCTCTGACAG 1200
DB 1141 TCTTGGAAATTAAGAACTCTCAAGAGTCTCTGCTGCTATTTCTGGCCCTCTGACAG 1200
QY 1201 GCTGTGACTCAGGATACCTGTGCTGCTCAAGTGAATCTCATTTGTTATTTGAGAA 1260
DB 1201 GCTGTGACTCAGGATACCTGTGCTGCTCAAGTGAATCTCATTTGTTATTTGAGAA 1260
QY 1261 GACAGGCTGCAAGGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCCGCTATTAC 1320
DB 1261 GACAGGCTGCAAGGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCCGCTATTAC 1320
QY 1321 TCTTCAGGCTCTTCTACTATTTGTTGTCACACAGCATCTCATGCTCTTCAAGGTTACT 1380
DB 1321 TCTTCAGGCTCTTCTACTATTTGTTGTCACACAGCATCTCATGCTCTTCAAGGTTACT 1380
QY 1381 CCAATGACTCCTTCTGCTCTTCAAGTGGAGCAATGCTTAAAGTGTATTAATTCATCT 1440
DB 1381 CCAATGACTCCTTCTGCTCTTCAAGTGGAGCAATGCTTAAAGTGTATTAATTCATCT 1440
QY 1441 ATTCTCTTGGCAATCTTCTTCTGAGGCTTACTATTTGCTTATTTCAAAAG 1500
DB 1441 ATTCTCTTGGCAATCTTCTTCTGAGGCTTACTATTTGCTTATTTCAAAAG 1500
QY 1501 CAATGTGCAAGAAAGAGTAAAGATGAATCAATTCATTTCCCTGTGGCT 1560

Db	1501	CAATGTCGCCAAGGAAAAGAAAGTTAAAGAAAGATGAAATTAATTCATTTCCTGGTGGCCT	1560
Qy	1561	GTGGGGGACTGTGTGCAAAACTACTCGTCTCCCTTTTACACGACTCCTTTGGCCAGAG	1620
Db	1561	GTGGGGGACTGTGTGCAAAACTACTCGTCTCCCTTTTACACGACTCCTTTGGCCAGAG	1620
Qy	1621	CAGAGAAATGAAAAAGCCAGGGGAGGTGGAANAATCGATGCTTCAGCTGTGGCTGTGCTGCC	1680
Db	1621	CAGAGAAATGAAAAAGCCAGGGGAGGTGGAANAATCGATGCTTCAGCTGTGGCTGTGCTGCC	1680
Qy	1681	AGCCAAAGTCTTCATTTGGGGCCAAAGGGGAAAACCTTTTGTGGAGAAAGCCCTGTGCTTT	1740
Db	1681	AGCCAAAGTCTTCATTTGGGGCCAAAGGGGAAAACCTTTTGTGGAGAAAGCCCTGTGCTTT	1740
Qy	1741	GTCAACCCAGCGCTGGAAATCAGTGGCGGGATCTCAGCTCACCGCAACCTTCACCTTCGGG	1800
Db	1741	GTCAACCCAGCGCTGGAAATCAGTGGCGGGATCTCAGCTCACCGCAACCTTCACCTTCGGG	1800
Qy	1801	TTCAAGTAAATTTCCCTGCTCAGGCTCCCAAGTATGCTGGAAATACAGGCACGCCACCATG	1860
Db	1801	TTCAAGTAAATTTCCCTGCTCAGGCTCCCAAGTATGCTGGAAATACAGGCACGCCACCATG	1860
Qy	1861	CCGAGCTAATTTTGTATTTTTCAGTAGAAACGGGATTCACACAGTGGCCAGGCTGTCTC	1920
Db	1861	CCGAGCTAATTTTGTATTTTTCAGTAGAAACGGGATTCACACAGTGGCCAGGCTGTCTC	1920
Qy	1921	TCGAATCTCTGACCGCAAGTATCCACCCGCTCCGCTCCCAAGTGTCTGGGATTAACAG	1980
Db	1921	TCGAATCTCTGACCGCAAGTATCCACCCGCTCCGCTCCCAAGTGTCTGGGATTAACAG	1980
Qy	1981	GCGTAGAGCAACCGTCCCGGCCCCGAAAGGGGAAACTCTGTGGAGAGACAGAGGGGCTCA	2040
Db	1981	GCGTAGAGCAACCGTCCCGGCCCCGAAAGGGGAAACTCTGTGGAGAGACAGAGGGGCTCA	2040
Qy	2041	CATCTCCCCCTGTGATTTCCCCCATGACATTTGCTTTATCTCTCCCATTTAGCCAGAAATC	2100
Db	2041	CATCTCCCCCTGTGATTTCCCCCATGACATTTGCTTTATCTCTCCCATTTAGCCAGAAATC	2100
Qy	2101	TATGTGTGTTTTCTTCTGTGCAATTTACTATGATTTGTATGATGTGCGGTACACACACCCCC	2160
Db	2101	TATGTGTGTTTTCTTCTGTGCAATTTACTATGATTTGTATGATGTGCGGTACACACACCCCC	2160
Qy	2161	CCCATGTGGGGGGGTGAGAGGGGGTGCAGAGCCCTGCTGCTCACCTTTTTCATCCTTGGAA	2220
Db	2161	CCCATGTGGGGGGGTGAGAGGGGGTGCAGAGCCCTGCTGCTCACCTTTTTCATCCTTGGAA	2220
Qy	2221	CTGATATTGATAAATCACTTCTGTGTTGTTCAGTTTTCAAAAAATTTTTTTTTT	2273
Db	2221	CTGATATTGATAAATCACTTCTGTGTTGTTCAGTTTTCAAAAAATTTTTTTTTT	2273

RESULT	
2	
AAf25169	
ID	AAf25169 standard; cDNA; 2264 BP.
XX	
AC	AAf25169;
XX	
XX	
D7	30-Apr-2001 (first entry)
DE	
Nucleotide sequence of a human protein having a hydrophobic domain.	
XX	
KW	Human; hydrophobic protein; secretory protein; membrane protein; sepsis;
KM	tumour inhibition; immune deficiency; autoimmune disorder; anaemia; burn
KM	infectious disease; cancer; ulcer; periodontal disease; coagulation;
KW	Parkinson's disease; fertility; immune response; thrombosis; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	85..1548
XX	/+tag= a
XX	/product= "hydrophobic protein"

PN	MO200104297-A2.
XX	18-JAN-2001.
PD	
XX	
PF	16-JUN-2000; 2000MO-JP03942.
PR	08-JUL-1999; 99JP-0194359.
XX	
PA	(SAGA) SAGAMI CHEM RES CENT.
PA	(PROT-) PROTEGENE INC.
PI	Kato S, Kimura T;
XX	
DR	WPI; 2001-103081/11.
DR	P-PSDB; AAB31669.
XX	
PT	Isolated human proteins and polynucleotides are used in research and
PT	have activities including cell proliferation/differentiation activity,
PT	immune stimulating activity and receptor/ligand activity -
XX	
PS	Claim 4; Page 120-124; 151pp; English.
XX	
CC	The present sequence encodes a human protein with hydrophobic domains.
CC	AA25159 represents a shorter version of the present sequence. The
CC	protein possesses a hydrophobic domain and so is a secretory protein
CC	or a membrane protein. The protein is used as an antigen to prepare
CC	antibodies. The polynucleotide sequence is useful as a source of probes
CC	for genetic diagnosis. It is also useful for producing the protein
CC	in large quantities and for gene therapy. The eukaryotic cells are used
CC	for detecting the receptors or ligands corresponding to the protein and
CC	for detecting small novel pharmaceuticals. The antibodies are also used
CC	for detection, quantification and purification of the proteins. Both the
CC	protein and polynucleotide may be used in research or as nutritional
CC	sources or supplements. The protein may have cytokine and cell
CC	proliferation/differentiation activity, immune stimulating or suppressing
CC	activity, hematopoiesis regulating activity, tissue growth activity,
CC	activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
CC	and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC	activity and tumour inhibition activity. It may therefore may be used to
CC	treat immune deficiencies resulting from autoimmune disorders or
CC	infectious diseases, cancer, sepsis, anemias, burns and ulcers.
CC	periodontal disease, Parkinson's disease, induce fertility, improve
CC	immune response and enhance coagulation or inhibit thrombosis.
XX	
SQ	Sequence 2264 BP; 482 A; 629 C; 550 G; 603 T; 0 other;
	Query Match 98.6%; Score 2241.2; DB 22; Length 2264;
	Best Local Similarity 99.8%; Pred. No. 0; Mismatches 3; Indels 1; Gaps 1
	Matches 2254; Conservative 0;
QY	1 GGGGGTGAAGCGAAT-CGTTTGGCCCGGATTCGGGGCCGCGGAGACTGGGGGGTCCCTGT 59
DB	7 GGGGGTGAAGCGAATGCGTTTGGCCCGGATTCGGGGCCGCGGAGACTGGGGGGTCCCTGT 66
QY	60 GGGGCTCCGGAGTTAAGATGGCGTCTCAGCGGAGGGGAGAGAGGAGACTGTGTGGCG 119
DB	67 GGGGCTCCCGGAGTTAAGATGGCGTCTCAGCGGAGGGGAGAGAGGAGACTGTGTGGCG 126
QY	120 CTGGCGGGGGGTTTCGAGTGGGGTTTCCAGAGCTGAGACCTTAACAATGGCCGAGCTCC 179
DB	127 CTGGCGGGGGGTTTCGAGTGGGGTTTCCAGAGCTGAGACCTTAACAATGGCCGAGCTCC 186
QY	180 CTGGCGCGGTGAGAAAGAGCGCTGCGGTGATCATCTCATCTTCTGGGTTAACCCCTTT 239
DB	187 CTGGCGCGGTGAGAAAGAGCGCTGCGGTGATCATCTCATCTTCTGGGTTAACCCCTTT 246
QY	240 GCTTGTGTTTATGGCATTAACCTTTTCTACAAGAGACCTAAGCTTCACTCACTCTTCAAT 299
DB	247 GCTTGTGTTTATGGCATTAACCTTTTCTACAAGAGACCTAAGCTTCACTCACTCTTCAAT 306
QY	300 ACCCTTACAGGCGCTCAATGCTATTTTAACTTTGAAACCAAGCTTACACATCCCTCG 359
DB	307 ACCCTTACAGGCGCTCAATGCTATTTTAACTTTGAAACCAAGCTTACACATCCCTCG 366

QY 360 CTGTGATTTGCTGCTAGTTCCTCATCTCTTGAATAGGCGCCGACCATGACTGCTGC 419
 DB 367 CTGTGATTTGCTGCTAGTTCCTCATCTCTTGAATAGGCGCCGACCATGACTGCTGC 426
 QY 420 CTGACTACCTTTGCTTCCAGATGCGTACCTCTGCGTGGATACATTAATCACTGCGACC 479
 DB 427 CTGACTACCTTTGCTTCCAGATGCGTACCTCTGCGTGGATACATTAATCACTGCGACC 486
 QY 480 GGCACATACGATATCAAGTGAAGATGCCAATGTGTTCTGACTTTGAAGCTGATTTGT 539
 DB 487 GGCACATACGATATCAAGTGAAGATGCCAATGTGTTCTGACTTTGAAGCTGATTTGT 546
 QY 540 TTGGCTGTTGACTACTTTTGAAGGAGAAAGATCAGAAATTCCTTGTCTCTGACCAAG 599
 DB 547 TTGGCTGTTGACTACTTTTGAAGGAGAAAGATCAGAAATTCCTTGTCTCTGACCAAG 606
 QY 600 AATATAGCAATACGATGCTGCTTCCCTGCTGGAAGTTGCTGCTTCTGCTATCTAT 659
 DB 607 AATATAGCAATACGATGCTGCTTCCCTGCTGGAAGTTGCTGCTTCTGCTATCTAT 666
 QY 660 GGGGCTCTTCTGTTAGGCGCCGACCTTCTCAATGATACATACATGAGCTGTGACAGGA 719
 DB 667 GGGGCTCTTCTGTTAGGCGCCGACCTTCTCAATGATACATACATGAGCTGTGACAGGA 726
 QY 720 GAGCTGATTTGACTATCCAGAAAGATACCAAGATCATCTCTGCTCTCAAGCGCTG 779
 DB 727 GAGCTGATTTGACTATCCAGAAAGATACCAAGATCATCTCTGCTCTCAAGCGCTG 786
 QY 780 AGTCTGAGGCTTTTCTACCTAGTGGGCTACACATGCTGACGCGCCGACATCAGAAAGAC 839
 DB 787 AGTCTGAGGCTTTTCTACCTAGTGGGCTACACATGCTGACGCGCCGACATCAGAAAGAC 846
 QY 840 TATCTCTCACTGAGACATATGACCAACACCCCTTCTGTTCCGCTGACATGATCATGCT 899
 DB 847 TATCTCTCACTGAGACATATGACCAACACCCCTTCTGTTCCGCTGACATGATCATGCT 906
 QY 900 ATCTGGGGCAAGTTTGTGCTGTACAAATATGTACCTGTTGGCTGTACAGAAAGATG 959
 DB 907 ATCTGGGGCAAGTTTGTGCTGTACAAATATGTACCTGTTGGCTGTACAGAAAGATG 966
 QY 960 TGCATTTTGAAGGCGCTTCAATGCTTGAAGAAAGGCAAGGCAAGTGGAGAT 1019
 DB 967 TGCATTTTGAAGGCGCTTCAATGCTTGAAGAAAGGCAAGGCAAGTGGAGAT 1026
 QY 1020 GCTGTGCAACATGAAAGGTGTGCTTTTGAACAAACCCCGCTTCACTGCGACCAT 1079
 DB 1027 GCTGTGCAACATGAAAGGTGTGCTTTTGAACAAACCCCGCTTCACTGCGACCAT 1086
 QY 1080 GCTCATTTCAACATCAACACCAACGCTGGTGGCGCTGACATCTTCAACAGATCAG 1139
 DB 1087 GCTCATTTCAACATCAACACCAACGCTGGTGGCGCTGACATCTTCAACAGATCAG 1146
 QY 1140 TTCTTTGAAATTAAGAACTCTCTCAGAGGTCTCTGTTGCTATCTGCGCTCTGAGAC 1199
 DB 1147 TTCTTTGAAATTAAGAACTCTCTCAGAGGTCTCTGTTGCTATCTGCGCTCTGAGAC 1206
 QY 1200 GGCCTGCACTCAGGATACCTGTGTCTTCCAGATGGAATTCCTCATTTGTTGGA 1259
 DB 1207 GGCCTGCACTCAGGATACCTGTGTCTTCCAGATGGAATTCCTCATTTGTTGGA 1266
 QY 1260 AGACAGGCTGGCGGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCGCCCATTTACT 1319
 DB 1267 AGACAGGCTGGCGGCTCATTTCAAGAGAGCCCACTGAGCAAGCTGGCGCCCATTTACT 1326
 QY 1320 GTCTCTCAGGCTTCTACTATTTGTTGACACAGACCATCAGCTGCTCTCATGGGTTAC 1379
 DB 1327 GTCTCTCAGGCTTCTACTATTTGTTGACACAGACCATCAGCTGCTCTCATGGGTTAC 1386
 QY 1380 TCCATGATGCTCTTGTGCTCTTCACTGTGGAGCAAAATGCTTAAAGTATTAATCCATC 1439
 DB 1387 TCCATGATGCTCTTGTGCTCTTCACTGTGGAGCAAAATGCTTAAAGTATTAATCCATC 1446

QY 1440 TATTTCTTTGGCCACATCTTTCTTCTGAGCTTACTATTAATTTGCTTAATTCAGAA 1499
 DB 1447 TATTTCTTTGGCCACATCTTTCTTCTGAGCTTACTATTAATTTGCTTAATTCAGAA 1506
 QY 1500 GCAATGTTGTCAGGAAAGAGAGATTAAGAAAGATGGAATTAATTCATTTCCCGTGGGCC 1559
 DB 1507 GCAATGTTGTCAGGAAAGAGAGATTAAGAAAGATGGAATTAATTCATTTCCCGTGGGCC 1566
 QY 1560 TGTGCGGAGCTGTGTCAGAAATCTCGTCTCCCTTTTACAGACATCTCTTTGCCAGAA 1619
 DB 1567 TGTGCGGAGCTGTGTCAGAAATCTCGTCTCCCTTTTACAGACATCTCTTTGCCAGAA 1626
 QY 1620 GCAGAGATTTGAAAGCCAGGAGGTGGAAGATCGATGCTTCCAGCTGTGCTCTGCTG 1679
 DB 1627 GCAGAGATTTGAAAGCCAGGAGGTGGAAGATCGATGCTTCCAGCTGTGCTCTGCTG 1686
 QY 1680 CAGCCAACTTCAATTTGGGGCCAAAGGGGAAATCTTTTGGAGAAAGCGCTTGTCT 1739
 DB 1687 CAGCCAACTTCAATTTGGGGCCAAAGGGGAAATCTTTTGGAGAAAGCGCTTGTCT 1746
 QY 1740 TGTACCCACGCTGGAATGCAATGCGGGATCTCAGCTCAGCCGCACTTCACTCTG 1799
 DB 1747 TGTACCCACGCTGGAATGCAATGCGGGATCTCAGCTCAGCCGCACTTCACTCTG 1806
 QY 1800 GTTCAAGTATTTTCTGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1859
 DB 1807 GTTCAAGTATTTTCTGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1866
 QY 1860 GCCCAGCTAATTTTGTATTTTCAATGAGAAAGGGAATTCACACAGCTTGGCAGGCTGT 1919
 DB 1867 GCCCAGCTAATTTTGTATTTTCAATGAGAAAGGGAATTCACACAGCTTGGCAGGCTGT 1926
 QY 1920 CTGGAATCTCTGACCGCAAGTATCCACCCGCTCGCTCCCAAGTCTGGGATTTACA 1979
 DB 1927 CTGGAATCTCTGACCGCAAGTATCCACCCGCTCGCTCCCAAGTCTGGGATTTACA 1986
 QY 1980 GGGGTGAGCCACCTGTCGCGCCGCAAGGGGAAATCTTGTGGAGAGACAGAGGGCTC 2039
 DB 1987 GGGGTGAGCCACCTGTCGCGCCGCAAGGGGAAATCTTGTGGAGAGACAGAGGGCTC 2046
 QY 2040 ACATCTCCCTCGATTTCCCATGACATGCTTATCTCTCCCATCTAGCCAGGAAT 2099
 DB 2047 ACATCTCCCTCGATTTCCCATGACATGCTTATCTCTCCCATCTAGCCAGGAAT 2106
 QY 2100 CTATGTGTTTCTTCTGCAATTTACTATGATGTGTATGTGCGGTACCAACACCCC 2159
 DB 2107 CTATGTGTTTCTTCTGCAATTTACTATGATGTGTATGTGCGGTACCAACACCCC 2166
 QY 2160 CCCCATGGGGGTGAGAGGGGTGCAAGGCCCTGCTGCTCAGCTTTTCTAACCTTTGA 2219
 DB 2167 CCCCATGGGGGTGAGAGGGGTGCAAGGCCCTGCTGCTCAGCTTTTCTAACCTTTGA 2226
 QY 2220 ACTGATTTAGATAAATCACTTCTGTTTGTGAGTTT 2257
 DB 2227 ACTGATTTAGATAAATCACTTCTGTTTGTGAGTTT 2264

RESULT 3
 AAZ50889
 ID AAZ50889 standard; cDNA; 2234 BP.
 XX
 AC AAZ50889;
 XX
 DE Human receptor-associated protein cDNA from Incyte clone 2906971.
 XX
 DT 31-MAY-2000 (first entry)
 XX
 KW Human receptor-associated protein; HRAP; Incyte clone 2906971;
 KW cytosolic; immunomodulatory; anti-inflammatory; cardiant; antianaemic;
 KW antiarteriosclerotic; hepatotropic; antiallergic antirheumatic;
 KW antistatic; osteopathic; antiallergic; antidiabetic; dermatological;
 KW neuroprotective; diagnosis; treatment; prevention; reproductive disorder;
 KW cardiovascular; cell proliferative; autoimmune; inflammatory; allergy;

Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	279	TACTCATCCACTCTTCCATACCTTACAGGCTCTCAATGCTATTTTAACCTTGA	338
Db	1	TACTCATCCACTCTTCCATACCTTACAGGCTCTCAATGCTATTTTAACCTTGA	60
Qy	339	AACGAGCTTACACTCCCTGTGTGTATTTGTCTTCAAGTCTCTCACTTCCGACTAATG	398
Db	61	AACGAGCTTACACTCCCTGTGTGTATTTGTCTTCAAGTCTCTCACTTCCGACTAATG	120
Qy	399	GGCCGACCATACAGTCCCTCTCTACTACTTTTGTCTTCAAGTCCGCTTCTGCT	458
Db	121	GGCCGACCATACAGTCCCTCTCTACTACTTTTGTCTTCAAGTCCGCTTCTGCT	180
Qy	459	GGATCATTTATACAGTCCGCACTACGATATCAAGTGAACAATGCACATTTGTT	518
Db	181	GGATCATTTATACAGTCCGCACTACGATATCAAGTGAACAATGCACATTTGTT	240
Qy	519	CTGACTTTGAAGCTGATTTGGCTGTGACTACTTTTGAAGGAGGAAAGTCAAGAT	578
Db	241	CTGACTTTGAAGCTGATTTGGCTGTGACTACTTTTGAAGGAGGAAAGTCAAGAT	300
Qy	579	TCCCTGTCTCTGAGGAGAAATATGCGATATCGTGTGCTTCTTCCCTGCTGAAGTT	638
Db	301	TCCCTGTCTCTGAGGAGAAATATGCGATATCGTGTGCTTCTTCCCTGCTGAAGTT	360
Qy	639	GCTGGTTTCTCTACTCTTATGAGGCTCTTGTGAGGAGGAGGCTTCAATGATCAT	698
Db	361	GCTGGTTTCTCTACTCTTATGAGGCTCTTGTGAGGAGGAGGCTTCAATGATCAT	420
Qy	699	TACATGAAGCTGTGAGGAGAGCTGATGATACATACAGAGAAAGTACCAACAGCATC	758
Db	421	TACATGAAGCTGTGAGGAGAGCTGATGATACATACAGAGAAAGTACCAACAGCATC	480
Qy	759	ATTCTGTCTCTCAAGGCTGTGAGTGGGCTTTTCTACTGTAGGGCTTACACATGCTC	818
Db	481	ATTCTGTCTCTCAAGGCTGTGAGTGGGCTTTTCTACTGTAGGGCTTACACATGCTC	540
Qy	819	AGCCCCCATCATCAGAGAACTATCTCATGAGAGATATGACAAACCCCTTCTGG	878
Db	541	AGCCCCCATCATCAGAGAACTATCTCATGAGAGATATGACAAACCCCTTCTGG	600
Qy	879	TTCGCTGATGATGATGCTGATCTGGGCAAGTTTGTCTGTACAAATATGTCACTGT	938
Db	601	TTCGCTGATGATGATGCTGATCTGGGCAAGTTTGTCTGTACAAATATGTCACTGT	660
Qy	939	TGGCTGTCTACAGAGAGATATGATTTTGAAGGCTGTGGCTTCAATGGCTTTGAAGA	998
Db	661	TGGCTGTCTACAGAGAGATATGATTTTGAAGGCTGTGGCTTCAATGGCTTTGAAGA	720
Qy	999	AAGGCAAGGCAAGTGGATGCTGTGCAATGAAAGTGTGGCTTTTGAACAAC	1058
Db	721	AAGGCAAGGCAAGTGGATGCTGTGCAATGAAAGTGTGGCTTTTGAACAAC	780
Qy	1059	CCCCGCTTCACTGAGCACTATGCTCATTAACATCAACACCAAGCCGTGGGCGCCG	1118
Db	781	CCCCGCTTCACTGAGCACTATGCTCATTAACATCAACACCAAGCCGTGGGCGCCG	840
Qy	1119	TACATCTTCAACAGACTCAAGTTCTTGGAAATTAAGAACTCTTCAAGGGTCTCTGTTG	1178
Db	841	TACATCTTCAACAGACTCAAGTTCTTGGAAATTAAGAACTCTTCAAGGGTCTCTGTTG	900
Qy	1179	CTATTCTGGGCTCTGAGAGGCTGCACTAGATATCTGTCTGCTTCCAGATGAA	1238
Db	901	CTATTCTGGGCTCTGAGAGGCTGCACTAGATATCTGTCTGCTTCCAGATGAA	960
Qy	1239	TTCCTCATTTATTTGTGAAAGACAGGCTGCAAGCTCATTTCAAGAGAGCCCACTTG	1298
Db	961	TTCCTCATTTATTTGTGAAAGACAGGCTGCAAGCTCATTTCAAGAGAGCCCACTTG	1020
Qy	1299	AGCAAGCTGGCGCCATTACTGTCTCCAGCCCTTCACTATTTGGTGAACAGACATC	1358
Db	1021	AGCAAGCTGGCGCCATTACTGTCTCCAGCCCTTCACTATTTGGTGAACAGACATC	1080

Qy	1359	CAGTGGCTCTTCATGGGTAACTCATGACTGGCTTCTGCTCTTCAAGGAGCAAAATG	1418
Db	1081	CAGTGGCTCTTCATGGGTAACTCATGACTGGCTTCTGCTCTTCAAGGAGCAAAATG	1140
Qy	1419	CTTAAGGTATTAATTCATATTTCTTGGCCACATCTTCTTCTGAGCTTACTATTC	1478
Db	1141	CTTAAGGTATTAATTCATATTTCTTGGCCACATCTTCTTCTGAGCTTACTATTC	1200
Qy	1479	ATATTGCTTATTTACAAAGCAATGGTCCCAAGAAAGAAAGTAAAGAAATGAA	1538
Db	1201	ATATTGCTTATTTACAAAGCAATGGTCCCAAGAAAGAAAGTAAAGAAATGAA	1260
Qy	1539	TATCCATTTCCCTGGTGGCTGTGCGGAACTGTGCAAGAACTGCTCCCTTTC	1598
Db	1261	TATCCATTTCCCTGGTGGCTGTGCGGAACTGTGCAAGAACTGCTCCCTTTC	1320
Qy	1599	ACAGCACTCTTTGCCCCAGAGCAGAGATGAAAGCCAGGAGTGAAGATCATGC	1658
Db	1321	ACAGCACTCTTTGCCCCAGAGCAGAGATGAAAGCCAGGAGTGAAGATCATGC	1380
Qy	1659	TTCAGCTGTGCTCTGCTGCCAGCAGTCTTCAATTTGGGGCCAAAGGGAACTTTT	1718
Db	1381	TTCAGCTGTGCTCTGCTGCCAGCAGTCTTCAATTTGGGGCCAAAGGGAACTTTT	1440
Qy	1719	TTTGAGAGAGGCTGCTTGTGACCCAGCAGTGAATGCAATGGCGGATCTGAGCTC	1778
Db	1441	TTTGAGAGAGGCTGCTTGTGACCCAGCAGTGAATGCAATGGCGGATCTGAGCTC	1500
Qy	1779	ACCGCACTCCACTCTCTGGGTTCAAGTATTTCTGCTCAAGCTTCCCAAGTAGCTG	1838
Db	1501	ACCGCACTCCACTCTCTGGGTTCAAGTATTTCTGCTCAAGCTTCCCAAGTAGCTG	1560
Qy	1839	GGAATACAGGCAAGCCACCATATGCCAGTAAATTTTGAATTTTCAATGAAACGGATTT	1898
Db	1561	GGAATACAGGCAAGCCACCATATGCCAGTAAATTTTGAATTTTCAATGAAACGGATTT	1620
Qy	1899	CACCAAGTGGCAGGCTGTGAGTCCCTGCAAGGAGATCCACCGGCTCCGCG	1958
Db	1621	CACCAAGTGGCAGGCTGTGAGTCCCTGCAAGGAGATCCACCGGCTCCGCG	1680
Qy	1959	TCCCAAGTGTGGATTACAGGCTGTGAGCCACCGTCCGCGCCCAAGGGGAACTTT	2018
Db	1681	TCCCAAGTGTGGATTACAGGCTGTGAGCCACCGTCCGCGCCCAAGGGGAACTTT	1740
Qy	2019	GTGGAGAGAGAGAGGGCTCATCTCCCTCTGATTTCCCATGCAATTCCTTATC	2078
Db	1741	GTGGAGAGAGAGAGGGCTCATCTCCCTCTGATTTCCCATGCAATTCCTTATC	1800
Qy	2079	TCTCCCATCTAGCAGGAATCTATTTGTTTCTTCTGCC	2120
Db	1801	TCTCCCATCTAGCAGGAATCTATTTGTTTCTTCTGCC	1842

RESULT 5
ABN96918
ID ABN96918 standard; DNA; 1842 BP.

ABN96918;
13-AUG-2002 (first entry)

Gene #3416 used to diagnose liver cancer.

Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.

Homo sapiens.

WO200229103-A2.

11-APR-2002.

Db 1561 GGAAATACAGCAGCAGCCACCATGCCAGTAATTTTGTATTTTCACTAGMAAGGGGATTT 1620
Qy 1899 CACCACTGTGGCAGGCTGATCTCGAATCTCTGACCGCAAGTATCCACCGCTCCGCC 1958
Db 1621 CACCACTGTGGCAGGCTGATCTCGAATCTCTGACCGCAAGTATCCACCGCTCCGCC 1680
Qy 1959 TCCCAAGTGTGGATTTACAGGCGTGAACCAACCGTCCCGCCAAAGGGAACTCTT 2018
Db 1681 TCCCAAGTGTGGATTTACAGGCGTGAACCAACCGTCCCGCCAAAGGGAACTCTT 1740
Qy 2019 GTGGAGAGCAGAGGGGCTCACATTTCCCTCTGATTTCCCGATGCACATTGCTTATC 2078
Db 1741 GTGGAGAGCAGAGGGGCTCACATTTCCCTCTGATTTCCCGATGCACATTGCTTATC 1800
Qy 2079 TCTCCCATCTAGCCAGCAATCATTTGTTTCTTCTGCC 2120
Db 1801 TCTCCCATCTAGCCAGCAATCATTTGTTTCTTCTGCC 1842

RESULT 6
ABL64098
ID ABL64098 standard; DNA; 1842 BP.

AC ABL64098;
XX
DT 15-MAY-2002 (first entry)

DE Breast cancer related gene sequence SEQ ID NO:2435.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytoskeletal; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; de.

OS Homo sapiens.

PN WO200194629-A2.

XX 13-DEC-2001.

PD 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209513P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.

PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;

DR WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -

XX Claim 1; SEQ ID 2435; 44pp; English.

XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (1) of a signature gene set, where (1)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (1) has cytoskeletal
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.

XX Sequence 1842 BP; 413 A; 525 C; 413 G; 491 T; 0 other;

XX Query Match 81.0%; Score 1842; DB 24; Length 1842;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 TACCTCATCCACCTCTCCATACCTTTACGCGCTCTCAATGCTTATTTTAACTTGA 338
Db 1 TACCTCATCCACCTCTCCATACCTTTACGCGCTCTCAATGCTTATTTTAACTTGA 60
Qy 339 AACCACTCTACCACTCCCTGCTGTGTATTTGCTTCACTTCATCTTGAATAAG 398
Db 61 AACCACTCTACCACTCCCTGCTGTGTATTTGCTTCACTTCATCTTGAATAAG 120
Qy 399 GGGCGCACCATCTGCGGTCTCACTACCTTTTCTTCAGATGGCTTACCTTGTGCT 458
Db 121 GGGCGCACCATCTGCGGTCTCACTACCTTTTCTTCAGATGGCTTACCTTGTGCT 180
Qy 459 GGATACCTTTACAGTCCGCAACCGCAATCAATCAAGTGGCAATGCATATGTGT 518
Db 181 GGATACCTTTACAGTCCGCAACCGCAATCAATCAAGTGGCAATGCATATGTGT 240
Qy 519 CTGACTTTGAAGCTGATTTGCTGTGCTTGAAGTGAAGGAGGAAGATCAGAT 578
Db 241 CTGACTTTGAAGCTGATTTGCTGTGCTTGAAGTGAAGGAGGAAGATCAGAT 300
Qy 579 TCTTGTCTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCTCTGGAAGTT 638
Db 301 TCTTGTCTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCTCTGGAAGTT 360

QY 639 GCTGTTCTCTACTTCTATGAGGSCCTTGTGTAGGGCCAGTTCATGAATCAC 638
 DB 361 GCTGTTCTCTACTTCTATGAGGSCCTTGTGTAGGGCCAGTTCATGAATCAC 420
 QY 699 TACATGAAGCTGTGTGAGGAGAGCTGATGACATACAGAAAGATACCAACAGATC 758
 DB 421 TACATGAAGCTGTGTGAGGAGAGCTGATGACATACAGAAAGATACCAACAGATC 480
 QY 759 ATTCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACTAGTGGCTACACATGCTC 818
 DB 481 ATTCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACTAGTGGCTACACATGCTC 540
 QY 819 AGCCCCCATCTCAGAGACTATCTCTCTGAAAGCTATGACCAACCCCTCTG 878
 DB 541 AGCCCCCATCTCAGAGACTATCTCTCTGAAAGCTATGACCAACCCCTCTG 600
 QY 879 TTCGCTGCTATGATGCTGATCTGGGCGAAGTTTGTCTGATCAAAATATGTCACCTG 938
 DB 601 TTCGCTGCTATGATGCTGATCTGGGCGAAGTTTGTCTGATCAAAATATGTCACCTG 660
 QY 939 TGGCTGTGTCAGAGAGAGTATGATTTTGAAGGCGCTGGGCTTCAATGGCTTTGAAGA 998
 DB 661 TGGCTGTGTCAGAGAGAGTATGATTTTGAAGGCGCTGGGCTTCAATGGCTTTGAAGA 720
 QY 999 AAGGGCAAGGCAAGTGGGATGCTGTGCCAATGAAGGTGGCTTTGAAGCAAC 1058
 DB 721 AAGGGCAAGGCAAGTGGGATGCTGTGCCAATGAAGGTGGCTTTGAAGCAAC 780
 QY 1059 CCCCCTTCATCTGGCAACATTCGCTCTCAATCAATCAACCAACGCTGGTGGCCGC 1118
 DB 781 CCCCCTTCATCTGGCAACATTCGCTCTCAATCAATCAACCAACGCTGGTGGCCGC 840
 QY 1119 TACATCTTCAAGAGCTCAAGTTCCTTGGAAATTAAGAACTCTCTCAGGGTCTCTGTTG 1178
 DB 841 TACATCTTCAAGAGAGCTCAAGTTCCTTGGAAATTAAGAACTCTCTCAGGGTCTCTGTTG 900
 QY 1179 CTAATCTGAGCCCTCTGGGCGAGGCTGCACTCAAGATACCTGGTCTGCTCCAGATGGA 1238
 DB 901 CTAATCTGAGCCCTCTGGGCGAGGCTGCACTCAAGATACCTGGTCTGCTCCAGATGGA 960
 QY 1239 TTCCTCATTTGTAATGTGAAAGACAGGCTGCCAGGCTCATTAAGAGAGCCGCTG 1298
 DB 961 TTCCTCATTTGTAATGTGAAAGACAGGCTGCCAGGCTCATTAAGAGAGCCGCTG 1020
 QY 1299 AGCAAGCTGGCCGCAATTAATGCTCTGACGCTTCTACTAATTTGGTGAACAGACCATC 1358
 DB 1021 AGCAAGCTGGCCGCAATTAATGCTCTGACGCTTCTACTAATTTGGTGAACAGACCATC 1080
 QY 1359 CACTGGCTCTTCANAGGCTTACTCATGATGCTTCTGCTTCACTGAGGGAACAAATG 1418
 DB 1081 CACTGGCTCTTCANAGGCTTACTCATGATGCTTCTGCTTCACTGAGGGAACAAATG 1140
 QY 1419 CTTAAGGTATTAATTCATCTATTTCTTGGCCACATCTTCTCTGAGCCTACTATTC 1478
 DB 1141 CTTAAGGTATTAATTCATCTATTTCTTGGCCACATCTTCTCTGAGCCTACTATTC 1200
 QY 1479 ATATTGCTTATTAATTCAGAAAGATGTGCAAGAAAGAAAGTAAAGATGGA 1538
 DB 1201 ATATTGCTTATTAATTCAGAAAGATGTGCAAGAAAGAAAGTAAAGATGGA 1260
 QY 1539 TAATTCATTTCCCTGGTGGCTGTGAGGGAAGTGTGAGAACTACTGCTCTCTTTC 1598
 DB 1261 TAATTCATTTCCCTGGTGGCTGTGAGGGAAGTGTGAGAACTACTGCTCTCTTTC 1320
 QY 1599 ACAGCACTCTTTGCCCCAGAGAGAGATGAAAGCCAGGAGGTGGAAGATGATGC 1658
 DB 1321 ACAGCACTCTTTGCCCCAGAGAGAGATGAAAGCCAGGAGGTGGAAGATGATGC 1380
 QY 1659 TTTCAGGTGTGCTCTGCTGCTGCAAGGCTTCTATTTGGGCGCAAGGGGAAACTTTT 1718
 DB 1381 TTTCAGGTGTGCTCTGCTGCTGCAAGGCTTCTATTTGGGCGCAAGGGGAAACTTTT 1440
 QY 1719 TTTCAGGAGAGGCGCTTGTGTTGTCAACCAAGCTGGAATGAGTGGGGATCTCAGCTC 1778

DB 1441 TTTGAGAAAGGCGCTCTTCTGTTGTACCCAGCTGGAATGCAATGCGGAGATTCAGCTC 1500
 QY 1779 ACCGCAACCTCCACCTCTGAGGTTCAAGTATTTTCTGCTTCAAGCTTCCAAAGTAC 1838
 DB 1501 ACCGCAACCTCCACCTCTGAGGTTCAAGTATTTTCTGCTTCAAGCTTCCAAAGTAC 1560
 QY 1839 GGAATACAGGACGCCACCAATGCGGCTAATTTTGTATTTTCAATAGAAAGGGAATTT 1898
 DB 1561 GGAATACAGGACGCCACCAATGCGGCTAATTTTGTATTTTCAATAGAAAGGGAATTT 1620
 QY 1899 CACCAAGTGGCCAGGCTGATCTGCAACTCTGCAACCGCAAGTATCCACCGGCTCGCC 1958
 DB 1621 CACCAAGTGGCCAGGCTGATCTGCAACTCTGCAACCGCAAGTATCCACCGGCTCGCC 1680
 QY 1959 TCCCAAGTGTGGATTAACAGCGTGAACCAACCGTGGCCGCAAGGGGAAACTTT 2018
 DB 1681 TCCCAAGTGTGGATTAACAGCGTGAACCAACCGTGGCCGCAAGGGGAAACTTT 1740
 QY 2019 GTGGAGAGAGCAGAGGGGCTCACATCTCCCTCTGATTTCCCATGACATGCTTATC 2078
 DB 1741 GTGGAGAGAGCAGAGGGGCTCACATCTCCCTCTGATTTCCCATGACATGCTTATC 1800
 QY 2079 TCTCCCATCTAGCCAGAACTATTTGTTTCTTCTGCC 2120
 DB 1801 TCTCCCATCTAGCCAGAACTATTTGTTTCTTCTGCC 1842

RESULT 7
 AAS70385
 ID AAS70385 standard; cDNA; 1867 BP.
 XX
 AC AAS70385;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #6189.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dermanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PDB; ABG06198.
 XX
 PT New isolated polynucleotide and encoded polypeptide, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID No 6189; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences, (I) is useful as hybridisation probe,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving

DB 1141 CTGGCTGCTTCAGATGGAATCCATGTTATATGGAAGAAGAGGCTGCGACAGCTC 1200
QY 1278 ATTCAGAGAGGCCCCCTGAGCAGCTGGCCGACATACGTCTCTCCAGCCCTTTCAC 1337
DB 1201 ATTCAGAGAGGCCCCCTGAGCAGCTGGCCGACATACGTCTCTCCAGCCCTTTCAC 1260
QY 1338 TATTGTGTCACAGACATCCATCTGCTCTTCAATGAGTTCATCCATGACTGCTTCTG 1397
DB 1261 TATTGTGTCACAGACATCCATCTGCTCTTCAATGAGTTCATCCATGACTGCTTCTG 1320
QY 1398 CTCTTCACCTGGAGCAATGCTTAAAGTGTATTAATCCATCTATTTCTTGCCACATC 1457
DB 1321 CTCTTCACCTGGAGCAATGCTTAAAGTGTATTAATCCATCTATTTCTTGCCACATC 1380
QY 1458 TCTTCCTGAGCCTCTATTCATATGCTTATATTCACAAGCAATGCTGCGAAGAA 1517
DB 1381 TCTTCCTGAGCCTCTATTCATATGCTTATATTCACAAGCAATGCTGCGAAGAA 1440
QY 1518 GAGAGTTAAGAGATGGA 1538
DB 1441 GAGAGTTAAGAGATGGA 1461

RESULT 9
ABV30301
ID ABV30301 standard; cDNA; 1296 BP.
XX

ABV30301;

16-SEP-2002 (first entry)

Human prostate expression marker cDNA 30292.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.

Homo sapiens.

MO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US05171.

17-FEB-2000; 2000US-183319P.

16-MAR-2000; 2000US-189862P.

25-MAY-2000; 2000US-207454P.

09-JUN-2000; 2000US-211314P.

18-JUL-2000; 2000US-219007P.

13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JB;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1, Page 6582; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 1296 BP; 324 A; 342 C; 320 G; 296 T; 14 other;

Query Match 34.0%; Score 773.6; DB 23; Length 1296;
Best Local Similarity 95.9%; Pred. No. 5.7e-206;
Matches 794; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 957 GTATGATTTTGAAGGCGCTGGGCTTCAATGCTTTGAAGAAAGGCAAGGCAAGTGG 1016
DB 91 GTATGATTTTGAAGGCGCTGGGCTTCAATGCTTTGAAGAAAGGCAAGGCAAGTGG 150
QY 1017 GATGCTGTGCGCAATGAAGGTGTGCTCTTTGAACAAACCCCGCTTCACTGGCAC 1076
DB 151 GATGCTGTGCGCAATGAAGGTGTGCTCTTTGAACAAACCCCGCTTCACTGGCAC 210
QY 1077 ATTGCTCTATTCATCAATCAACCAAGCGCTGGGTGGCCCGCTACATCTTCAAGCATC 1136
DB 211 ATGCGCTCATTCATCAATCAACCAAGCGCTGGGTGGCCCGCTACATCTTCAAGCATC 270
QY 1137 AAGTCTTGGAATTAAGAACTCTCTGAGGCTCTGCTGCTATTCCTGCGCCCTTGG 1196
DB 271 AAGTCTTGGAATTAAGAACTCTCTGAGGCTCTGCTGCTATTCCTGCGCCCTTGG 330
QY 1197 CAGGCGCTGCACTCAGATACCTGTGCTCTTCCAGATGAATTCCTCATTTGTTG 1256
DB 331 CAGGCGCTGCACTCAGATACCTGTGCTCTTCCAGATGAATTCCTCATTTGTTG 390
QY 1257 GAAAGCAGGCTGCGCAGGCTCATTCAGAGAGCCCCCAGCTGAGCAAGCTGCGCCATT 1316
DB 391 GAAAGCAGGCTGCGCAGGCTCATTCAGAGAGCCCCCAGCTGAGCAAGCTGCGCCATT 450
QY 1317 ACTGCTCTGAGGCGCTTCTCAATTTGTGTCAGACAGACCATCAGCTGCTTCAATGGT 1376
DB 451 ACTGCTCTGAGGCGCTTCTCAATTTGTGTCAGACAGACCATCAGCTGCTTCAATGGT 510
QY 1377 TACTCCATGACTGCTTCTGCTCTTCACTGTCAGTGGCAAAATGCTTAAGGTATTAATCC 1436
DB 511 TACTCCATGACTGCTTCTGCTCTTCACTGTCAGTGGCAAAATGCTTAAGGTATTAATCC 570
QY 1437 ATCTATTTCTTGGCCACATCTTCTTCTGAGCTTCAATTCATATGCTTATATTCAC 1496
DB 571 ATCTATTTCTTGGCCACATCTTCTTCTGAGCTTCAATTCATATGCTTATATTCAC 630
QY 1497 AAGCAATGTCGCCAAGAAAGAGAAAGTAAAGATGGAATATTCATTTCCCTGTCG 1556
DB 631 AAGCAATGTCGCCAAGAAAGAGAAAGTAAAGATGGAATATTCATTTCCCTGTCG 690
QY 1557 GCTGTGGGGAGTGTGTCAGAACTACTGCTCTCTTTTTCACAGCACTCTTTGGCCCC 1616
DB 691 GCTGTGGGGAGTGTGTCAGAACTACTGCTCTCTTTTTCACAGCACTCTTTGGCCCC 750
QY 1617 AGAGCAGAGATGGAAGAAAGCCAGGAGGTGAAGATGATGCTTCAAGTGTGCTTGC 1676
DB 751 AGAGCAGAGATGGAAGAAAGCCAGGAGGTGAAGATGATGCTTCAAGTGTGCTTGC 810
QY 1677 TGCCAGCAAGTCTCATTTTGGGGCCAAAGGGGAATCTTTTGGAGAGAGCGCTTGG 1736
DB 811 TGCCAGCAAGTCTCATTTTGGGGCCAAAGGGGAATCTTTTGGAGAGAGCGCTTGC 870
QY 1737 CTCTGTCAACCAAGCTGGAATGATGAGTGGCGGATCTCAGCTACCGCA 1784
DB 871 TTGTGTACCAACAGCTGGAATGCACTGCGGGATCTCAGCTACCGGA 918

RESULT 10
AAS30639/C
ID AAS30639 standard; DNA; 7461 BP.

PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0234001.
PR 14-SEP-2000; 2000US-023401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-476224/51.
The present invention relates to the isolation of novel human
respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
sequences encoding for these polypeptides. The sequences of the
invention are useful for preventing, treating and/or prognosing
disorders related to the respiratory system including respiratory
disorders related to the respiratory system including throat

CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
CC of the invention are useful in gene therapy and antisense therapy.
CC AAS28161-AAS28764 represent genomic sequences encoding for novel
CC human respiratory antigens.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 7461 BP; 1883 A; 1666 C; 1923 G; 1969 T; 0 other;
Query Match 31.2%; Score 708.6; DB 22; Length 7461;
Best Local Similarity 99.4%; Pred. No. 2,3e-187;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1546 TTCCCTGGTGGCTGTGGCGGAGCTGGTGAGAACTACTGCTCCCTTTTCACAGCAC 1605
DB 721 TCTTGCAGGTGGCTGTGGCGGAGCTGGTGAGAACTACTGCTCCCTTTTCACAGCAC 662
QY 1606 TCCCTTGGCCCCAGAGCAGAGAAATGAAAAAGCCAGGGAGGTGAGAAATGATGCTTCACG 1665
DB 661 TCCCTTGGCCCCAGAGCAGAGAAATGAAAAAGCCAGGGAGGTGAGAAATGATGCTTCACG 602
QY 1666 TGTGCTCTGTCTGCTCCAGCCAGCTTCATTTGGGGCCAAAGGGGAAACTTTTGGAG 1725
DB 601 TGTGCTCTGTCTGCTCCAGCCAGCTTCATTTGGGGCCAAAGGGGAAACTTTTGGAG 542
QY 1726 AAGGCGTCTGCTTTGTTCACCCAGCTGGAATGAGTGGGGGATCTCAGCTCACCGCAA 1785
DB 541 AAGGCGTCTGCTTTGTTCACCCAGCTGGAATGAGTGGGGGATCTCAGCTCACCGCAA 482
QY 1786 CCTCACTCTCTGGGTTCAAGTATTTCTGCTCAGCTCCCAAGTAGCTGGGAATAC 1845
DB 481 CCTCACTCTCTGGGTTCAAGTATTTCTGCTCAGCTCCCAAGTAGCTGGGAATAC 422
QY 1846 AGGCAAGCCCACTGCCCACTAATTTTGTATTTTCAATGAGAAAGGATTTCCACGAC 1905
DB 421 AGGCAAGCCCACTGCCCACTAATTTTGTATTTTCAATGAGAAAGGATTTCCACGAC 362
QY 1906 TTGGCCAGGCTGTCTGTAAGTCTCTGACCCGAGTAGTCCACCGCTCCGCCCAAA 1965
DB 361 TTGGCCAGGCTGTCTGTAAGTCTCTGACCCGAGTAGTCCACCGCTCCGCCCAAA 302
QY 1966 GTGCTGGGATTACAGGGGTGAGCCACCGTCCCGGCCAAAGGGGAAACTCTGTGGAG 2025
DB 301 GTGCTGGGATTACAGGGGTGAGCCACCGTCCCGGCCAAAGGGGAAACTCTGTGGAG 242
QY 2026 GAGCAGAGGGGCTCAATCTCCCTCGATTCCGCCATGACATGCTGCTTATCTTCCCC 2085
DB 241 GAGCAGAGGGGCTCAATCTCCCTCGATTCCGCCATGACATGCTGCTTATCTTCCCC 182
QY 2086 ATCTAGCAGGAATCTATTGTGTTTCTTCTGCAATTTACTATGATGTGATGACC 2145
DB 181 ATCTAGCAGGAATCTATTGTGTTTCTTCTGCAATTTACTATGATGTGATGACC 122
QY 2146 GCTACCAACACCCCCCATGGGGGGGTGAGAGGGGGTGAAGGCCCTGCTCTCCACT 2205
DB 121 GCTACCAACACCCCCCATGGGGGGGTGAGAGGGGGTGAAGGCCCTGCTCTCCACT 62
QY 2206 TTTTCTAAGCTTGAAGTGTATGATTAATCACTTGTGTTGTTCAGTTTTCAC 2260
DB 61 TTTTCTAAGCTTGAAGTGTATGATTAATCACTTGTGTTGTTCAGTTTTCAC 7
RESULT 12
ID ACA03402/c
AC ACA03402; standard; DNA; 7461 BP.
XX
XX
DT 22-MAY-2003 (first entry)

XX DNA encoding human lung cancer antigen HIPA05.
DE
XX
XX Human; ds; gene; lung cancer antigen; anaemia; tissue regeneration;
KW constant region; immunotherapy; gene therapy; lung disorder; leukopenia;
KW small cell lung cancer; squamous cell carcinoma; adenocarcinoma; asthma;
KW respiratory disorder; nonallergic rhinitis; rheumatoid arthritis; ulcer;
KW adult respiratory distress syndrome; hyperproliferative disorder; graft;
KW lung neoplasm; prostate neoplasm; immune system disorder; anaphylaxis;
KW multiple sclerosis; blood-related disorder; allergic reaction; impotence;
KW inflammatory disorder; appendicitis; dermatitis; graft vs host disease;
KW immune complex disease; serum sickness; polyarteritis nodosa; arhythmia;
KW urinary system disorder; glomerulonephritis; kidney failure; infertility;
KW cardiovascular disorder; myocardial infarction; osteoarthritis; aging;
KW musculoskeletal system disorder; Albers-Schönberg disease; cancer; burn;
KW neurological disorder; Alzheimer's disease; Parkinson's disease; trauma;
KW endocrine disorder; Addison's disease; diabetes mellitus; gastric reflux;
KW gastrointestinal disorder; reproductive system disorder; cerebral palsy;
KW developmental disorder; Fanconi's syndrome; cellular level disease;
KW amyotrophic lateral sclerosis; infectious disease; viral infection;
KW bacterial infection; wound healing.
XX
XX Homo sapiens.
OS
XX
XX US2002173454-A1.
PN
XX
XX 21-NOV-2002.
PD
XX
XX 17-JAN-2001; 2001US-0764904.
PF
XX
XX 31-JUN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220963P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.

PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBEN/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 WPI; 2003-311001/30.
 XX
 DR Novel lung cancer related polypeptide useful for detecting, treating,
 XX preventing and/or prognosing lung, respiratory, hyperproliferative,
 PT immune system, cardiovascular endocrine or gastrointestinal disorders
 PT
 PS Disclosure; Page 282-286; 332pp; English.
 XX
 CC The invention relates to an isolated lung cancer related polypeptide its
 CC variants, allelic variants or homologues. The polypeptide and its nucleic
 CC acid are useful for preventing, treating, or ameliorating a medical
 CC condition in a mammalian subject, for diagnosing a pathological condition
 CC or a susceptibility to a pathological condition in a subject by
 CC determining the presence or absence of mutation in the nucleic acid or
 CC biological sample and diagnosing a pathological condition based on the
 CC result. The polypeptide, antibodies to the polypeptide or the
 CC polynucleotide are useful for detecting, treating, preventing and/or
 CC prognosing disorders of the lung, such as small cell lung cancer, non-
 CC small cell lung cancer (e.g. squamous cell carcinoma, adenocarcinoma,
 CC large cell carcinoma, adenocarcinoma and undifferentiated
 CC carcinoma) or lung cancer metastasis. The polypeptide, its antibodies or
 CC its polynucleotide are also useful for detecting, treating, preventing
 CC and/or prognosing respiratory disorders e.g. nonallergic rhinitis and
 CC adult respiratory distress syndrome; hyperproliferative disorders e.g.
 CC lung neoplasms and prostate neoplasms; immune system disorders e.g.
 CC rheumatoid arthritis and multiple sclerosis; blood-related disorders e.g.
 CC anaemia and leukopenia; allergic reactions e.g. anaphylaxis and asthma;
 CC inflammatory disorders e.g. appendicitis and dermatitis; graft vs host
 CC disease; immune complex disease e.g. serum sickness and polyarteritis
 CC nodosa; urinary system disorders e.g. glomerulonephritis and kidney
 CC failure; cardiovascular disorders e.g. arrhythmia and myocardial
 CC infarction; musculoskeletal system disorders e.g. Albers-Schonberg
 CC disease and osteoarthritis; neurological disorders e.g. Alzheimer's
 CC disease and Parkinson's disease; endocrine disorders e.g. Addison's
 CC disease and diabetes mellitus; gastrointestinal disorders e.g. ulcers and
 CC gastric reflux; reproductive system disorders e.g. infertility and
 CC impotence; developmental and inherited disorders e.g. Fanconi's syndrome
 CC and cerebral palsy; diseases at the cellular level e.g. cancer and
 CC amyotrophic lateral sclerosis; infectious diseases e.g. viral and
 CC bacterial. The polypeptide, its antibodies or its polynucleotide are also
 CC useful for wound healing e.g. grafts and burns and for regeneration of
 CC tissues damaged due to e.g. aging and trauma. The present sequence
 CC represents DNA encoding a human lung cancer antigen.
 CC
 XX
 SQ Sequence 7461 BP; 1883 A; 1666 C; 1923 G; 1989 T; 0 other;

Query Match 31.2%; Score 708.6; DB 25; Length 7461;
 Best Local Similarity 99.4%; Pred. No. 2.3e-187;
 Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1546 TTTCCTGTGCTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 1605
 DB 721 TCTTGACAGTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 662
 QY 1606 TCTTTTGTGCTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 1665
 DB 661 TCTTTTGTGCTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 602

QY 1666 TGTGCTGTGCTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 1725
 DB 601 TGTGCTGTGCTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 542
 QY 1726 AAGGCGCTGTGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 1785
 DB 541 AAGGCGCTGTGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 482
 QY 1786 CCTGCACTCTGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 1845
 DB 481 CCTGCACTCTGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 422
 QY 1846 AGGCAAGCAAGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 1905
 DB 421 AGGCAAGCAAGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 362
 QY 1906 TTGGCAGAGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 1965
 DB 361 TTGGCAGAGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 302
 QY 1966 GTGCTGTGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 2025
 DB 301 GTGCTGTGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 242
 QY 2026 GAGCAGAGAGGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 2085
 DB 241 GAGCAGAGAGGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 182
 QY 2086 ATCTGAGCAGAGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 2145
 DB 181 ATCTGAGCAGAGCTGTGCTGTGCGAGCTGTGACAGAACTACTCTCTCCCTTTACAGCAGC 122
 QY 2146 GCTAC 2205
 DB 121 GCTAC 62
 QY 2206 TTTTCTACTCTGTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2260
 DB 61 TTTTCTACTCTGTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7
 RESULT 13
 ID ABA15665/3
 ID ABA15665 standard; DNA; 32174 BP.
 AC ABA15665;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 7996.
 XX
 KW Human; noctropic; neuroprotective; cytostatic; dermatological; virocidic;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antischizy; antianaemic; antiaerthetic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antinflammatory;
 KW antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.

PR	17-MAR-2000	2000US-0198023
PR	18-APR-2000	2000US-0191075
PR	19-MAY-2000	2000US-0205515
PR	07-JUN-2000	2000US-0209467
PR	28-JUN-2000	2000US-0214886
PR	30-JUN-2000	2000US-0215535
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216680
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225566
PR	14-AUG-2000	2000US-0225567
PR	18-AUG-2000	2000US-0225679
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227182
PR	30-AUG-2000	2000US-0227009
PR	01-SEP-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	12-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234274
PR	25-SEP-2000	2000US-0234997
PR	25-SEP-2000	2000US-0234998
PR	25-SEP-2000	2000US-0235484
PR	26-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235634
PR	27-SEP-2000	2000US-0235635
PR	29-SEP-2000	2000US-0236127
PR	29-SEP-2000	2000US-0236127
PR	29-SEP-2000	2000US-0236167
PR	29-SEP-2000	2000US-0236168
PR	29-SEP-2000	2000US-0236598
PR	29-SEP-2000	2000US-0236599
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935

PR	13-OCT-2000;	2000US-0239937.
ER	20-OCT-2000;	2000US-0240360.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	20-OCT-2000;	2000US-0242221.
PR	01-NOV-2000;	2000US-0246117.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250391.
PR	01-DEC-2000;	2000US-0251160.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251038.
PR	05-DEC-2000;	2000US-0251186.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251893.
PR	08-DEC-2000;	2000US-0251890.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Baraesh SC, Ruben SM,	
XX	WPI; 2001-541565/60.	
DR		
XX		
PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating nervous system	
PT	cancers and metastases	
XX		
XX		
PS	Disclosure; SEQ ID NO 7996; 1701tp + Sequence Listing; English.	
XX		
CC	The invention relates to novel genes (ABA11004-ABA21534) and proteins	
CC	(ABA1678-ABA18001) useful for preventing and/or ameliorating	

QY 1846 AGGACGCCACCATGCTCCAGCTAATTTTGTATTTTCAAGTAGAAGGATTTCCACAG 1905
DB 26425 AGGACGCCACCATGCTCCAGCTAATTTTGTATTTTCAAGTAGAAGGATTTCCACAG 26484
QY 1906 TTGGCCAGGCTGGTCTCGAACTCTGACCGCAAGTATCCACCCGCTCCGCTCCCAA 1965
DB 26485 TTGGCCAGGCTGGTCTCGAACTCTGACCGCAAGTATCCACCCGCTCCGCTCCCAA 26544
QY 1966 GTGCTGGATTACAGGCGTAGCCACGTCGCGCCGCAAGGGGAACTTTGTGGAG 2025
DB 26545 GTGCTGGATTACAGGCGTAGCCACGTCGCGCCGCAAGGGGAACTTTGTGGAG 26604
QY 2026 GAGCAGAGGGGCTCACATCTCCCTCTGATTCCCGCATGCATTCCTTATCTCCCG 2085
DB 26605 GAGCAGAGGGGCTCACATCTCCCTCTGATTCCCGCATGCATTCCTTATCTCCCG 26664
QY 2086 ATCTAGCCAGATCTATTGTGTTTCTTCTGCAATTTTACATATGTGATGTGCC 2145
DB 26665 ATCTAGCCAGATCTATTGTGTTTCTTCTGCAATTTTACATATGTGATGTGCC 26724
QY 2146 GCTACCAACACCCCGCCATGGGGGGGAGAGAGGGGTCGAAGGCCCTGCTGCTCACT 2205
DB 26725 GCTACCAACACCCCGCCATGGGGGGGAGAGAGGGGTCGAAGGCCCTGCTGCTCACT 26784
QY 2206 TTTTCTACCTTGAATCTGATTAGATMAATCATCTCTGTTTCTGATTTTCA 2260
DB 26785 TTTTCTACCTTGAATCTGATTAGATMAATCATCTCTGTTTCTGATTTTCA 26839

RESULT 15
ID ABA20359 standard; DNA; 32174 BP.
XX ABA20359;
AC
XX
XX 23-JAN-2002 (first entry)
XX
XX
DE Human nervous system related polynucleotide SEQ ID NO 12690.
XX
XX
KW Human; nocotropic; neuroprotective; cyrostatic; dermatological; virucide;
KW immunosuppressive; anti-inflamatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antistickling; antianaemic; antiathrictic; cancer;
KW antilemmatic; hepatotropic; cerebroprotective; antinflammation;
KW antiallergic; antidiabetic; antilicer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0232404.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI: 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Disclosure; SEQ ID NO 12690; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABAI1004-ABAI21534) and proteins
 CC (ABAI4678-ABAI5001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 32174 BP; 8747 A; 7064 C; 6837 G; 8001 T; 1525 other;

Query Match 31.2%; Score 708.6; DB 22; Length 32174;
 Best Local Similarity 99.4%; Pred. No. 5.1e-187;
 Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1546	TTTCCCTGGTGGGCTGTGGGGAAGTGTGCAGAAACTACTGCTCTCTTTTCACAGCAC	1605
DB	26125	TCCTGAGGTGGCTGTGGGGAAGTGTGCAGAAACTACTGCTCTCTTTTCACAGCAC	26184
QY	1606	TCCTTTGCCCCAGAGCAGAAATGAAAAAGCCAGGAGGTGGAATGATGCTTCACG	1665
DB	26185	TCCTTTGCCCCAGAGCAGAAATGAAAAAGCCAGGAGGTGGAATGATGCTTCACG	26244
QY	1666	TGTGCTCTGTGCTGCCAGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGAG	1725
DB	26245	TGTGCTCTGTGCTGCCAGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGAG	26304
QY	1726	AAGGCGCTTGTGCTTTGTCAACCCAGCTGGAATGACAGTGGCGGGATCTCAGCTCACCGCA	1785
DB	26305	AAGGCGCTTGTGCTTTGTCAACCCAGCTGGAATGACAGTGGCGGGATCTCAGCTCACCGCA	26364
QY	1786	CCTGCACCTCTGTGGGTTCAAGTATTTCTGCTTCAGCTCCCAAGTACTGGGAATAC	1845
DB	26365	CCTGCACCTCTGTGGGTTCAAGTATTTCTGCTTCAGCTCCCAAGTACTGGGAATAC	26424
QY	1846	AGGCAAGCCACCATGCGCCAGCAATTTTGTATTTTCACTAATAAAGGGATTTCAACAGG	1905
DB	26425	AGGCAAGCCACCATGCGCCAGCAATTTTGTATTTTCACTAATAAAGGGATTTCAACAGG	26484
QY	1906	TTGGCAGAGCTGTGCTCGAATCTCGAACCGCAAGTATCACCGGCTCCGCCCA	1965
DB	26485	TTGGCAGAGCTGTGCTCGAATCTCGAACCGCAAGTATCACCGGCTCCGCCCA	26544
QY	1966	GTGCTGGATTAACAGCGGTGAGCCACCGTCCCGGCCAAAGGGGAACTTTGTGGAG	2025
DB	26545	GTGCTGGATTAACAGCGGTGAGCCACCGTCCCGGCCAAAGGGGAACTTTGTGGAG	26604
QY	2026	GAGCAGAGGGGCTCATCTCCCTCTGATTTCCCAATGACATGCTTATCTCTCC	2085
DB	26605	GAGCAGAGGGGCTCATCTCCCTCTGATTTCCCAATGACATGCTTATCTCTCC	26664
QY	2086	ATCTAGCCAGAAATCTATTGTTTTCTTCTGCCAATTTACTATGATGTGATGTC	2145
DB	26665	ATCTAGCCAGAAATCTATTGTTTTCTTCTGCCAATTTACTATGATGTGATGTC	26724
QY	2146	GCTACCAACACCCCTCCCATGCGGGGGTGAAGAGGGGTGCAAGGCTCTGCTCACT	2205
DB	26725	GCTACCAACACCCCTCCCATGCGGGGGTGAAGAGGGGTGCAAGGCTCTGCTCACT	26784
QY	2206	TTTTTACCTTGAATCTGATTAATGAATCACTCTGTTGTTCAGTTTCA	2260
DB	26785	TTTTTACCTTGAATCTGATTAATGAATCACTCTGTTGTTCAGTTTCA	26839

Search completed: November 22, 2003, 02:18:40
 Job time: 605 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2003, 02:06:23 ; Search time 4740 Seconds
(without alignments)

11554.868 Million cell updates/sec

Title: US-09-938-803-26

Perfect score: 2273

Sequence: 1 99999tgaagcagatagctt.....ttttcaaaaaaaaaaaaaa 2273

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inh:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_phg:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1036.2	45.6	1201	13	BX402797 BX402797
2	1002	44.1	1130	9	AL550083 AL550083
3	952.4	41.9	1754	11	AK083687 Mus muscu
4	930.4	40.9	1000	9	AL532090 AL532090

5	907.6	39.9	1149	13	BX339313
6	906.4	39.9	1053	12	BM557200
7	906.2	39.9	958	9	AL532089
8	873.4	38.4	1201	9	AL557774
9	864.8	38.0	1065	9	AL574064
10	856.2	37.7	1201	9	AL579737
11	855.2	37.6	1201	9	AL576170
12	843.2	37.1	1097	12	BM555135
13	837.6	36.8	976	13	BX431852
14	836.4	36.8	974	13	BQ917856
15	828.4	36.4	882	12	B1767794
16	822.6	36.2	887	10	BG753831
17	818.2	36.0	911	14	CA489207
18	814.4	35.8	929	10	BG740629
19	806.6	35.5	916	13	BUS41837
20	800.2	35.2	905	13	BUS56427
21	794.4	34.9	847	12	B1760522
22	793.2	34.9	897	9	AU149574
23	789.8	34.7	1201	9	AL553069
24	789.4	34.7	1252	10	BG684877
25	777.8	34.2	800	12	BG823157
26	776.6	34.2	967	10	BG696449
27	772.4	34.0	893	13	BUS53426
28	765.4	33.7	884	12	B1199710
29	763.6	33.6	827	13	BUS97281
30	761	33.5	994	13	B0675002
31	758.8	33.4	791	9	AU126035
32	757.8	33.3	815	9	AU124635
33	757.6	33.3	858	10	BG757592
34	751.8	33.1	1069	13	BQ067563
35	751.6	33.1	947	13	BUS55890
36	750.8	33.0	787	9	AU130894
37	749.8	33.0	830	9	AU133043
38	744.8	32.8	835	9	AU143336
39	738.2	32.5	1134	12	BG824361
40	735.6	32.4	798	9	AU131310
41	733	32.2	744	13	BX098372
42	721	31.7	787	9	BG758511
43	715.2	31.5	787	9	AU125489
44	711.6	31.3	816	10	BG695903
45	709.8	31.2	737	12	BG772149

ALIGNMENTS

RESULT 1
BX402797/c
LOCUS
DEFINITION BX402797 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens CDNA clone CSDDU003YA18 3-PRIME, mRNA sequence.
ACCESSION BX402797
VERSION BX402797.1 GI:30622908
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6809.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSIAJ0012E10NP1cluster=6809.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS1AJ001ZE10NP1.

FEATURES
source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ003YA18"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 290 a 246 c 355 g 265 t 45 others

ORIGIN

Query Match 45 6%; Score 1036.2; DB 13; Length 1201;
Best Local Similarity 92.3%; Pred. No. 5.3e-16;
Matches 1096; Conservative 37; Mismatches 47; Indels 7; Gaps 5;

1031 CATGAAGTGTGGCTTTGAACAACCCCGCTTCACTGCACCATTCCTCATTCAA 1090
1181 CAACATGAGGTGTTTGAACAAACCCCGCTTCACTGCACCATTCCTCATTCAA 1124
1091 CATCAACACCAACGCTGGTGGCGGCTTCACTGCACCATTCCTCATTCAA 1150
1123 CATCAACACCAACGCTGGTGGCGGCTTCACTGCACCATTCCTCATTCAA 1066
1151 TAAAGAACTCTCTCAGGGTCTCGTTGCTATCTCGGGCTTGGACGGCTGCATC 1210
1065 TAAAGAACTCTCTCAGGGTCTCGTTGCTATCTCGGGCTTGGACGGCTGCATC 1006
1211 AGGATACCTGTGCTGCTTCAGATGGA--TTCTCATTTGTTATGGAAGACAGGCTG 1269
1005 AGGATACCTGTGCTGCTTCAGATGGA--TTCTCATTTGTTATGGAAGACAGGCTG 946
1270 CCAAGCTCATTCAGAGAGCCCAACCTGACAGAGCTGGCGGCATTAATCTCTCCAGC 1329
945 CCAAGCTCATTCAGAGAGCCCAACCTGACAGAGCTGGCGGCATTAATCTCTCCAGC 886
1330 CCTTACTATTGTTGGCAACAGACCATTCAGGGCTTATGAGGTTATTCATGACAGT 1389
885 CCTTACTATTGTTGGCAACAGACCATTCAGGGCTTATGAGGTTATTCATGACAGT 826
1390 CCTTACTATTGTTGGCAACAGACCATTCAGGGCTTATGAGGTTATTCATGACAGT 1449
825 CCTTACTATTGTTGGCAACAGACCATTCAGGGCTTATGAGGTTATTCATGACAGT 766
1450 GCCACATCTTCTCTGAGCTTATTCATTTGCTTATTCATTCACAAAGCATGCTG 1509
765 GCCACATCTTCTCTGAGCTTATTCATTTGCTTATTCATTCACAAAGCATGCTG 706
1510 CAAAGAAAGAGATTAAAGAAATGGAATATTCATTTCCCTGGTGGCGGCAAT 1569
705 CAAAGAAAGAGATTAAAGAAATGGAATATTCATTTCCCTGGTGGCGGCAAT 646
1570 TGTGTGAGAACTACTGCTCTCCCTTTTCAACAGACTCTTTGCCCAAGAGAGATG 1629
645 TGTGTGAGAACTACTGCTCTCCCTTTTCAACAGACTCTTTGCCCAAGAGAGATG 586
1630 GAAAAGCAGGAGGAGTGAAGATGATGCTTCCAGCTGCTCTGCTGCAGCAAGTC 1689
585 GAAAAGCAGGAGGAGTGAAGATGATGCTTCCAGCTGCTCTGCTGCAGCAAGTC 526
1690 TTGATTTGGGCAAAAGGAACTTTTGGAGAGGAGCTTGGCTTGTGACCCAC 1749
525 TTGATTTGGGCAAAAGGAACTTTTGGAGAGGAGCTTGGCTTGTGACCCAC 466
1750 GCTGAATGAGTGGCGGATCTCAGCTCAACGCACTCCTCTGCTGATCAAGTGA 1809
465 GCTGAATGAGTGGCGGATCTCAGCTCAACGCACTCCTCTGCTGATCAAGTGA 406

1810 TTTTCTGCTCAGCTCCCAAGTAGTGGGAATACAGGACGCCACCATGCTCACTAA 1869
405 TTTTCTGCTCAGCTCCCAAGTAGTGGGAATACAGGACGCCACCATGCTCACTAA 346
1870 TTTTGTATTTTTCAGTAGAAGGAGATTTCACAGTTTGGCGAGCTGCTCGAATCTC 1929
345 TTTTGTATTTTTCAGTAGAAGGAGATTTCACAGTTTGGCGAGCTGCTCGAATCTC 286
1930 TGAACGCAAGTATCCACCCGCTTCGCTCCCAAGTCTGGATTACAGCGGTGAGCT 1989
285 TGAACGCAAGTATCCACCCGCTTCGCTCCCAAGTCTGGATTACAGCGGTGAGCT 226
1990 ACCGTGCGCGCGCCCAAGGAGAACTCTTGTGTGGAGAGACCAAGGGCTCATATCTCCC 2049
225 ACCGTGCGCGCGCCCAAGGAGAACTCTTGTGTGGAGAGACCAAGGGCTCATATCTCCC 166
2050 TGTGATTTCCCATGACATTCCTTATCTCTCCCATCTAGCCAGAAATCTATTGTGT 2109
165 TGTGATTTCCCATGACATTCCTTATCTCTCCCATCTAGCCAGAAATCTATTGTGT 106
2110 TTTTCTGCTCAGTATTCATGATTTGATGTTGATGCGGCTACCAACGCCCGCCATGAGG 2169
105 TTTTCTGCTCAGTATTCATGATTTGATGTTGATGCGGCTACCAACGCCCGCCATGAGG 46
2170 GGGTGAAGAGGAGTGAAGGCGCTGCTGCTCATCTTTTCTACCTT 2216
45 SSTGCGMCCCGCCGCAAGCGCCCGCC--CTCCACATCAATCACT 1

RESULT 2
AL550083 1130 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION
AL550083 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI040YC05 5-PRIME, mRNA sequence.
AL550083
AL550083.2 GI:31271901
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1130)
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12886703.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6809.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI040AB03QPLcluster=6809.f. Contact:
Feng Liang Email: fliang@life.techn.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DI040AB03QPL.
FEATURES
source
1. .1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI040YC05"
/issue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 240 a 292 c 294 g 299 t 5 others

Query Match 44.1%; Score 1002; DB 9; Length 1130;
 Best Local Similarity 98.4%; Pred. No. 3.2e-15;
 Matches 1038; Conservative 4; Mismatches 10; Indels 3; Gaps 3;

1 GGGGGTGAAGCAGTATGCTTTGCGGCGATTCGGGGGCGCGGAGCTGGGGGGTCCCTGTG 60
 DB GGGGGTGAAGCAGTATGCTTTGCGGCGATTCGGGGGCGCGGAGCTGGGGGGTCCCTGTG 137
 61 GGGGCTCCGGAGTAAAGATGGGCTCTCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 DB GGGGCTCCGGAGTAAAGATGGGCTCTCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 197
 121 TGGCGGGGGTCTGAGTGGGGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 DB TGGCGGGGGTCTGAGTGGGGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 257
 181 TGGGGGCGGTGAGAGCAGGCGGTGCGGCTGATCATCTTCATCTTCTGGGTTACCCCTTTG 240
 DB TGGGGGCGGTGAGAGCAGGCGGTGCGGCTGATCATCTTCATCTTCTGGGTTACCCCTTTG 317
 241 CTTGTTTATGGGAGTACCTTTCTAAGAGAGCCTCATCTCCACCTCTTCCATA 300
 DB CTTGTTTATGGGAGTACCTTTCTAAGAGAGCCTCATCTCCACCTCTTCCATA 377
 301 CCTTTACAGGCGCTTCATTTGCTTATTTTAACTTTGGAACAGGCTTACCACTCCCTGC 360
 DB CCTTTACAGGCGCTTCATTTGCTTATTTTAACTTTGGAACAGGCTTACCACTCCCTGC 437
 378 CTTTACAGGCGCTTCATTTGCTTATTTTAACTTTGGAACAGGCTTACCACTCCCTGC 437
 481 GCACTACAGTATCAAGTGAAGATGCAATGCTGCTGAGCTTTGAGAGTGGTGT 540
 DB GCACTACAGTATCAAGTGAAGATGCAATGCTGCTGAGCTTTGAGAGTGGTGT 617
 541 TGGCTGTGACTACTTTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
 DB TGGCTGTGACTACTTTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 677
 601 AATATGCCATAGTGTCTCTCCCTGCTGGAAGTCTGCTGCTTCTCTTCTATG 660
 DB AATATGCCATAGTGTCTCTCCCTGCTGGAAGTCTGCTGCTTCTCTTCTATG 737
 678 AATATGCCATAGTGTCTCTCCCTGCTGGAAGTCTGCTGCTTCTCTTCTATG 737
 661 GGGGCTTTGGTGAAGGCGCCAGTCTCAATGAATCACTACATGAGTGTGACAGGAGG 720
 DB GGGGCTTTGGTGAAGGCGCCAGTCTCAATGAATCACTACATGAGTGTGACAGGAGG 797
 738 GGGGCTTTGGTGAAGGCGCCAGTCTCAATGAATCACTACATGAGTGTGACAGGAGG 797
 721 AGCTGATGACATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 DB AGCTGATGACATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857
 788 AGCTGATGACATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857
 781 GTCCTGGGCTTTTAACTAGTGGGTACACATGCTGACGCGCCCATACAGAGAGT 840
 DB GTCCTGGGCTTTTAACTAGTGGGTACACATGCTGACGCGCCCATACAGAGAGT 917
 858 GTCCTGGGCTTTTAACTAGTGGGTACACATGCTGACGCGCCCATACAGAGAGT 917
 841 ATCTCTCACTAAGATGACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
 DB ATCTCTCACTAAGATGACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 977
 918 ATCTCTCACTAAGATGACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 977
 901 TCTGGGCGAGTTTGTCTGTAACAATATGCACTGTTGGTGTGACAGAGAGAT 960
 DB TCTGGGCGAGTTTGTCTGTAACAATATGCACTGTTGGTGTGACAGAGAGAT 1037
 978 TCTGGGCGAGTTTGTCTGTAACAATATGCACTGTTGGTGTGACAGAGAGAT 1037
 961 GGAATTTGACGGGCGGCTTCAATGGCTTTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1020
 DB GGAATTTGACGGGCGGCTTCAATGGCTTTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1096
 1038 GGAATTTGACGGGCGGCTTCAATGGCTTTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1096

Qy 1021 CCTGTGCCACATGAGGTGTGCTCTTTGAACA 1055
 Db 1097 CCTGTGCC-AMATGAGGTGTGCTCTTTGAAMAA 1130

RESULT 3
 AK083687
 LOCUS
 DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone: D03066012 product: unknown EST, full insert
 sequence.

ACCESSION AK083687
 VERSION AK083687.1 GI:26350712
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation

JOURNAL
REFERENCE
AUTHORS
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1754)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imclai, K., Ishii, Y., Itoh, S., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Takami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Murakami, M., and Hayashizaki, Y.

TITLE
JOURNAL
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifier

FEATURES

source

1..1754
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:D030066012"
/db_xref="taxon:10090"
/clone="D030066012"
/clone_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9 days embryo"
<1..1125
/note="unnamed protein product; putative
unknown EST (GB|BF302972, evidence: BLASTN, 96%,
match=736)"
/codon_start=1
/protein_id="BAC38993.1"
/db_xref="GI:26350713"
/translation="VITTLCPQMAVLLAGVYATGVDYDKMTMPHCVLTLKLGIC
DYDGGKDNLSLTSBOOKYAIRGVPSLLVAGSVPYGAFLVQPSMHHMYKLVQ
LNDIPKMRNSTIPALKRLSLGLVYVGTLLSPHITDYLITEDLNRPFRCKYH
LWGRFVLKYVTCWLVTEGVCLISGLNGEDENGSTVMDACANKWLVETTPFN
GTIASFNINTMNAVARYIFKRLFLKNGEISGLSLFLALMGHLSGLVLFQMEFL
IVYVERQVSLIRDSPLASISITLQFPYVVOOTIIMFLFEGYSMTAFCLTMDKW
LKYRSIYFLGHVFLSLFLLPIYIKHAWPRKRXKKE"
1731..1736
/note="putative"

CDS

polya_signal
/note="putative"
1754
polya_site
/note="putative"
BASE COUNT 437 a 477 c 383 g 457 t
ORIGIN

Query Match 41.9%; Score 952.4; DB 11; Length 1754;
Best Local Similarity 85.4%; Pred. No. 2.9e-14;
Matches 1090; Conservative 0; Mismatches 171; Indels 15; Gaps 2;

QY 417 GTCCTACTACCTTTGCTTCAGATGGCTACTCTTGCTGGCTGGAATATTACACTGCG 476
DB 1 GTTATTACTACCTTTGCTTCAGATGGCTACTCTTGCTGGCTGGAATATTACTACAGCC 60
QY 477 ACCGCACTACATATCAAGTGAGCAATGCAATGTGTCTGACTTTGAAGCTATT 536
||||| ||||||| ||||||| ||||||| ||| |||||||

DB 61 ACCGGTACTACATATCAAGTGAGCAATGCCAATTTGTCTTGACACTGAAGCTATT 120
QY 537 GCTTGGCTGTGACTACTCTTTGACGAGGGAAGATCAGAAATTCCTGTCTGTGACAA 596
DB 121 GGGCTGTGATTGACTACTACATGAGGCAAGAACGGGAATTCCTGTGACAA 180
QY 597 CAGAAATATGCAATACGTGTCTTCCCTGCTGGAAGTTGCGGTTTCTCTTCTTC 656
DB 181 CAGAAATATGCAATACGTGTCTTCCCTGCTGGAAGTTGCGGTTTCTCTTCTTC 240
QY 657 TATGGGCTTGTGTGTAAGGCGCCAGTTCTCAATGAATCACTACATGAAGTGTGACG 716
DB 241 TATGAGCTTCTTGTGTGTAAGGCGCCAGTTTCAATGAATCACTACATGAAGTGTG 300
QY 717 GGAAGCTATGACATACAGAAAGATACCAACAGCATATCTCTCTCAAGCCG 776
DB 301 GGAAGCTATGACATACAGAAAGATACCAACAGCATATCTCTCTCAAGCCG 360
QY 777 CTGAGCTGTGGGCTTTTCTACCTAGTGGGCTACACCTGTGAGCCGCAATACAGAA 836
DB 361 CTGAGCTGTGGGCTTTTCTACCTAGTGGGCTACACCTGTGAGCCGCAATACAGAA 420
QY 837 GACTATCTCTCACTGAAAGACTATGACCAACACCCCTTCTGCTCGCTGATATCATG 896
DB 421 GACTATCTCTCACTGAAAGACTATGACCAACACCCCTTCTGCTCGCTGATATCATG 480
QY 897 CTGATCTGGGCAAGTTTGTCTGTACAAATATGTCACCTGTGGTGTGACAGAA 956
DB 481 CTGATCTGGGCAAGTTTGTCTGTACAAATATGTCACCTGTGGTGTGACAGAA 540
QY 957 GTATGCAATTTTACGGGCTGTGGGCTTCAATGCTTTGAAGAAAGGCAAGCAAGTGT 1016
DB 541 GTGTGCTATCTGTGGGCTGTGGGCTTCAATGCTTTGAAGAAAGGCAAGTGT 600
QY 1017 GATGCTGTGCAACATGAAGTGTGCTCTTTGAAGAAAGGCAAGGCTTCACTGTGAC 1076
DB 601 GATGCTGTGCAACATGAAGTGTGCTCTTTGAAGAAAGGCAAGGCTTCACTGTGAC 660
QY 1077 ATTGCTCTTCAACATCAACACCAAGGCTGTGGGCTGTACATCTTCAAGACATC 1136
DB 661 ATGCTCTTCAACATCAACACCAAGGCTGTGGGCTGTACATCTTCAAGACATC 720
QY 1137 AAGTCTCTGAAATTAAGAACTCTCTGAGGCTCTGCTGTATTCCTGTGACCTCTG 1196
DB 721 AAGTCTCTGAAATTAAGAACTCTCTGAGGCTCTGCTGTATTCCTGTGACCTCTG 780
QY 1197 CAGGCTGCTCACTGAGATACCTGTGCTCTTCAAGATGAATTCCTATTTATGTG 1256
DB 781 CATGGCTCACTGAGATACCTGTGCTCTTCAAGATGAATTCCTATTTATGTG 840
QY 1257 GAAAGACAGGCTGCAAGCTCATTTCAAGAGACCCCACTTGAAGCAAGTGTGCGCAT 1316
DB 841 GAAAGACAGGCTGCAAGCTCATTTCAAGAGACCCCACTTGAAGCAAGTGTGCGCAT 900
QY 1317 ACTGTCCTCAAGCCCTTCACTATTGTGGCAAGACATCACTGTGCTTCAATGGT 1376
DB 901 ACTGTCCTCAAGCCCTTCACTATTGTGGCAAGACATCACTGTGCTTCAATGGT 960
QY 1377 TACTCATGACTGCTTCTGCTCTTCACTGCTGGAACAAATGCTTAAAGTGTATAATTC 1436
DB 961 TACTCATGACTGCTTCTGCTCTTCACTGCTGGAACAAATGCTTAAAGTGTATAATTC 1020
QY 1437 ATCTATTTCTTGGCCACATCTTTCTGAGACCTCAATTAATTTGCTTATATTCAC 1496
DB 1021 ATCTATTTCTTGGCCACATCTTTCTGAGACCTCAATTAATTTGCTTATATTCAC 1080
QY 1497 AAAGCAATGCTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1551
DB 1081 AAAGCAATGCTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1140
QY 1552 TGTGTGCTGTGCGGAGCTGTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1611
DB 1141 TGTGTGCTGTGCGGAGCTGTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1200

QY 1612 GCCCAGAGCAGAGATGAGAAAGCCAGGAGTGAAGATGATGCTTCAGCTGTGCC 1671
DB 1201 CCC-----AGCAGAGAGCCAGAGCTGGAAGTCAAGACACTCTCCAGCTGTGCC 1250
QY 1672 TCTGCTCCAGCCAG 1687
DB 1251 TCTGCGACAGCCCAAG 1266

RESULT 4
AL532090 1000 bp mRNA linear EST 23-MAY-2003
LOCUS AL532090 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM004YH02 5-PRIME, mRNA sequence.
ACCESSION AL532090
VERSION AL532090.2 GI:31069922
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1000)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12795583.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6809.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM004YH02&cluster=6809.f. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DM004YH02.1QPI.
location/Qualifiers
1.1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004YH02"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 210 a 266 c 247 g 272 t 5 others
ORIGIN

Query Match 40.9%; Score 930.4; DB 9; Length 1000;
Best Local Similarity 98.5%; Pred.No. 1.4e-13;
Matches 938; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

QY 22 GCCCGCATTCGGGGGCGCGGAGCTGAGGGGGTCCCTGTGGGGCTCCCGAGTTAAGATG 81
DB 49 GTCGGGAATTCGGGGATCGAGAGCTGGGGGTCCTCTGTGGGCTCCCGAGTTAAGATG 108
QY 82 GCTCTCAGCGGAGGGGAGCGAGGAGCTGTGTGGCGCTGGCGGGGTTCTGCAGTCGG 141
DB 109 GCTCTCAGCGGAGGGGAGCGAGGAGCTGTGTGGCGCTGGCGGGGTTCTGCAGTCGG 168
QY 142 GTTTCAGAGCTGAGCTTAACTTAAGTTGGGAGCGTCCCTGGCGCGGTGAGAAAGCGGC 201
DB 169 GTTTCAGAGCTGAGCTTAACTTAAGTTGGGAGCGTCCCTGGCGCGGTGAGAAAGCGGC 228
QY 202 TGGCGTATCATCTTCATCTCTGTGGTTACCCCTTGTCTTTTATCGGCAATTAC 261

DB 229 TGGCGTATCATCTTCATCTCTGTGGTTACCCCTTGTCTTTTATCGGCAATTAC 288
QY 262 TTTTCTAAGAGAACCTCACTCACTCACTCTTTCATACCTTTAAGAGCTCTCAATTG 321
DB 289 TTTTCTAAGAGAACCTCACTCACTCACTCTTTCATACCTTTAAGAGCTCTCAATTG 348
QY 322 CTTATTTTAACTTTGGAAGACAGCTCACTCCCTCTGTGATTTGCTTCAAGTCC 381
DB 349 CTTATTTTAACTTTGGAAGACAGCTCACTCCCTCTGTGATTTGCTTCAAGTCC 408
QY 382 TCATCTTTCGACTAATGAGCGGACATCACTGCGCTCTCACTACCTTTTCTTCAGA 441
DB 409 TCATCTTTCGACTAATGAGCGGACATCACTGCGCTCTCACTACCTTTTCTTCAGA 468
QY 442 TGGCTTACTCTTGTGCTGATTAATTAATCACTGCGGCACTAATCAAGTGA 501
DB 469 TGGCTTACTCTTGTGCTGATTAATTAATCACTGCGGCACTAATCAAGTGA 528
QY 502 CAATGCCAATTTGTTCTGACTTTGAAGCTGATTTGAGCTGTGACTTACTTGAAG 561
DB 529 CAATGCCAATTTGTTCTGACTTTGAAGCTGATTTGAGCTGTGACTTACTTGAAG 588
QY 562 GAGGGAAGATCAGAAATTCCTTGTCTGAGCAACAGAAATATGCAATAGCTGTTC 621
DB 589 GAGGGAAGATCAGAAATTCCTTGTCTGAGCAACAGAAATATGCAATAGCTGTTC 648
QY 622 CTTCCTCTCTGGAAGTTGCTGTCTTCTTCACTTGTAGGGCCCTTTGAGGGCCCC 681
DB 649 CTTCCTCTCTGGAAGTTGCTGTCTTCTTCACTTGTAGGGCCCTTTGAGGGCCCC 708
QY 682 AGTTCTCAATGATCACTACATGAGAGCTGAGGAGAGCTGATTAATCAACAGGA 741
DB 709 AGTTCTCAATGATCACTACATGAGAGCTGAGGAGAGCTGATTAATCAACAGGA 768
QY 742 AGATACCAAGACATCACTTCTGCTCAAGCGCTGAGTGTGGCCCTTTTCACTAG 801
DB 769 AGATACCAAGACATCACTTCTGCTCAAGCGCTGAGTGTGGCCCTTTTCACTAG 828
QY 802 TGGGTACACACTGTCTGCGCCCACTACAGAAAGATTAATCTCTCACTGAAGCTATG 861
DB 829 TGGGTACACACTGTCTGCGCCCACTACAGAAAGATTAATCTCTCACTGAAGCTATG 888
QY 862 ACAACCAACCCCTTGTGCTGCTGATGATGATGAGGGAAGTTTGCTGT 921
DB 889 ACAACCAACCCCTTGTGCTGCTGATGATGATGAGGGAAGTTTGCTGT 948
QY 922 ACAATATGATCACTGTGTGCTGCTGATGATGATGAGGGAAGTTTGCTGT 973
DB 949 ACAATATGATCACTGTGTGCTGCTGATGATGATGAGGGAAGTTTGCTGT 1000

RESULT 5
BX339313 1149 bp mRNA linear EST 02-MAY-2003
LOCUS BX339313 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS01070Y016 5-PRIME, mRNA sequence.
ACCESSION BX339313
VERSION BX339313.1 GI:30335787
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1149)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 6809.f. For more information about this cluster, see

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1070DE080P1&cluster=6809.f>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1070DE080P1.

FEATURES

source

1. 1149

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1070XJ16"

/issue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT

240 a 285 c 285 g 298 t 41 others

ORIGIN

Query Match 39.9%; Score 907.6; DB 13; Length 1149;
Best Local Similarity 97.9%; Pred. No. 3.9e-13;
Matches 924; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

23 CCCGATTCGGGGCGCGGAGCTGGGGGCTCTGTGGGGCTCCCGAGTTAAGATGCG 82
53 CCGGAATTCGGGGATCGCGAGCTGGGGGCTCTGTGGGGCTCCCGAGTTAAGATGCG 112
83 GTTCCTAGGAGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 142
113 GTTCCTAGGAGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 172
143 TTTCAGAGAGCTGAGCTTACAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAG 202
173 TTTCAGAGAGCTGAGCTTACAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAG 232
203 GGGGCTGATCATCTCCATCTTCTCTGGGTTACCCCTTGTCTTTTATGCGATTACCT 262
233 GGGGCTGATCATCTCCATCTTCTCTGGGTTACCCCTTGTCTTTTATGCGATTACCT 292
263 TTTCAGAGAGAGCTGAGCTTACAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAG 322
293 TTTCAGAGAGAGCTGAGCTTACAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAG 352
323 TTATTTAATTGGAAGCAGCTTACAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAG 382
353 TTATTTAATTGGAAGCAGCTTACAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAG 412
383 CATCTTGAATGAGGCGGAGCAGTCACTGCGGCTCTCACTACCTTTTGTCTTCAGAT 442
413 CATCTTGAATGAGGCGGAGCAGTCACTGCGGCTCTCACTACCTTTTGTCTTCAGAT 472
443 GGGCTACCTTGGCGGAGTACATTAAGTCACTGCGGAGCAGTCACTGCGGAGCAGTCA 502
473 GGGCTACCTTGGCGGAGTACATTAAGTCACTGCGGAGCAGTCACTGCGGAGCAGTCA 532
503 AATGACCATGTTGTTCTGACTTTGAAGTGAATGTTGGCTGTGACTACTTTCAGCG 562
533 AATGACCATGTTGTTCTGACTTTGAAGTGAATGTTGGCTGTGACTACTTTCAGCG 592
563 AGGGAAGATCAGAAATTCCTTGTCTGAGCAAGAAATATGCAATGCTGTGCTTC 622
593 AGGGAAGATCAGAAATTCCTTGTCTGAGCAAGAAATATGCAATGCTGTGCTTC 652
623 TTCCCTGAGGAGTGGTGTCTCTCACTTCAAGGAGCCTTCTGTGAGGAGCCCA 682
653 TTCCCTGAGGAGTGGTGTCTCTCACTTCAAGGAGCCTTCTGTGAGGAGCCCA 712
683 GTTCTCAATGATCCTACATGAGAGTGTGAGGAGAGCTGATTGACATACAGAGAA 742
713 GTTCTCAATGATCCTACATGAGAGTGTGAGGAGAGCTGATTGACATACAGAGAA 772

QY 743 GATACCAACAGACATCATCTTCTGCTCTCAAGCGCCTGAGTGGGCTTTTCTACTACT 802
DB 773 GATACCAACAGACATCATCTTCTGCTCTCAAGCGCCTGAGTGGGCTTTTCTACTACT 832
QY 803 GGGCTACACAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 862
DB 833 GGGCTACACAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 892
QY 863 CAACACCCCTTCTGTTCCGCTGAGTATGATGCTGAGGAGGAGGAGGAGGAGGAGGAG 922
DB 893 CAACACCCCTTCTGTTCCGCTGAGTATGATGCTGAGGAGGAGGAGGAGGAGGAGGAG 952
QY 923 CAATATGTCACCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
DB 953 CAATATGTCACCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 995

RESULT 6

BM557200

LOCUS

DEFINITION

AGENCOURT 6579051 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466619

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: DCTD/DTMP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LHCMI967 row: e column: 20

High quality sequence stop: 677.

FEATURES

source

1. 1053

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5466619"

/tissue_type="amelanotic melanoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 41"

/note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library"

BASE COUNT

230 a 313 c 237 g 267 t 6 others

ORIGIN

Query Match 39.9%; Score 906.4; DB 12; Length 1053;
Best Local Similarity 95.4%; Pred. No. 4.4e-13;
Matches 962; Conservative 0; Mismatches 42; Indels 4; Gaps 3;

QY 1092 ATCAACCAACGCGCTGGGTCGCGCTACATCTTCAACGACTCAAGTTCTTGAAT 1151
DB 1 ATCAACCAACGCGCTGGGTCGCGCTACATCTTCAACGACTCAAGTTCTTGAAT 60

QY	1152	AAAGAACTCTCTCAGAGGCTCTCGTTGGTAATTCCTGGGCGCTGTGGCAGGGCGCTGCAC	121
Db	61	AAAGAACTCTCTCAGAGGCTCTCGTTGGTAATTCCTGGGCGCTGTGGCAGGGCGCTGCAC	120
QY	1212	GGATACCTGCTCTGCTTCAGATGGAATTCCTCATTTGTAATGTGAAAGAAGACAGCTGCC	1272
Db	121	GGATACCTGCTCTGCTTCAGATGGAATTCCTCATTTGTAATGTGAAAGAAGACAGCTGCC	180
QY	1272	AGGCTCATTTCAAGAGAGGCCCACTGTAGCAAGCTGGCCGCGCATTACTGTCTCCAGCCC	1331
Db	181	AGGCTCATTTCAAGAGAGGCCCACTGTAGCAAGCTGGCCGCGCATTACTGTCTCCAGCCC	240
QY	1332	TTCTACATTTTGGTGCACAAGCCATCACTGGGCTTCATGGGTACTCCATGACTGCC	1397
Db	241	TTCTACATTTTGGTGCACAAGCCATCACTGGGCTTCATGGGTACTCCATGACTGCC	300
QY	1392	TTCTGCCCTCTTCACGCTGGGACAAATGGCTTAAGGTATTAATCATCATATTTCTTGAC	1457
Db	301	TTCTGCCCTCTTCACGCTGGGACAAATGGCTTAAGGTATTAATCATCATATTTCTTGAC	360
QY	1452	CACATCTTCTTCGTGAGCTTACTAATTCATATTGCCCTTAATTACAAAGCAATGGTGCCA	1511
Db	361	CACATCTTCTTCGTGAGCTTACTAATTCATATTGCCCTTAATTACAAAGCAATGGTGCCA	420
QY	1512	AGGAAGAGAAAGTTAAAGAAGTGAATAATTCATTTCCGCTGGGCGCTGGCCGGGACTG	1577
Db	421	AGGAAGAGAAAGTTAAAGAAGTGAATAATTCATTTCCGCTGGGCGCTGGCCGGGACTG	480
QY	1572	GTGCGAAGAACTACTGCTCTCCCTTTTTCACAGCACTCTTTGGCCGAGAGCAGAAATGGA	1631
Db	481	GTGCGAAGAACTACTGCTCTCCCTTTTTCACAGCACTCTTTGGCCGAGAGCAGAAATGGA	540
QY	1632	AAAGCCAGGAGAGGTGGAAGATCGATGCTTCAGCTGTGCTCTGTGTCAGCCAGAGTCTT	1691
Db	541	AAAGCCAGGAGAGGTGGAAGATCGATGCTTCAGCTGTGCTCTGTGTCAGCCAGAGTCTT	600
QY	1692	CATTTGGGGCCAAAGGGGAAACTTTTTTGGAGAAAGCGCTTTCCTTTGTCAACCCAGCG	1751
Db	601	CATTTGGGGCCAAAGGGGAAACTTTTTTGGAGAAAGCGCTTTCCTTTGTCAACCCAGCG	660
QY	1752	TGGAATGAGTGGCGGGATCTCAGCTCAACCGCACTCCACTCTGGGTTCAATGATTT	1811
Db	661	TGGAATGAGTGGCGGGATCTCAGCTCAACCGCACTCCACTCTGGGTTCAATGATTT	720
QY	1812	TTCTGTGCTCAGCTCCCAAGTAGCTGGGAATCAGGCAAGCCACATGATGCCAGTAAT	1871
Db	721	TTCTGTGCTCAGCTCCCAAGTAGCTGGGAATCAGGCAAGCCACATGATGCCAGTAAT	780
QY	1872	TTTGTATTTTTCAGTAGAAGACGGGATTTTCAACACGTTTGGCCAGGCTGTCTGAATCTCTG	1933
Db	781	TTTGTATTTTTCAGTAGAAGACGGGATTTTCAACACGTTTGGCCAGGCTGTCTGAATCTCTG	840
QY	1932	ACCAGAAATGATTCACACCGGCTCCGCTCCCAATGTGTGGGATTACA--GGGCTGAGCCA	1992
Db	841	ACCAGCAATGATTCACACCGGCTCCGCTCCCAATGTGTGGGATTACAAGGGCTGAGCCA	900
QY	1991	CCGTGCCCGGCCCAAGGGGAAACTTTTG--TGAGAGAGCAGAGGGGCTCACTCT--CC	2047
Db	901	CCGTGCCCGGCCCAAGGGGAAACTTTTGTGGAGAGNACAAAGGGGCTCACTCTCC	960
QY	2048	CCTCTGATTTCCCCCATGCACTGTGCTTATCTCTCCCATCTAGCCAG	2095
Db	961	CCTCTGATTTCCCCCATGCACTTTGCCCTTATCTCTCCCATCTAG	1008

RESULT 7	AL532089/c	958 bp	mRNA	linear	EST 23-MAY-2003
LOCUS	AL532089				
DEFINITION	AL532089 Homo sapiens FETAL LIVER Homo sapiens CDNA clone				
ACCESSION	CSDDM004YH02.3-PRIME, mRNA sequence.				
VERSION	AL532089				
KEYWORDS	AL532089.2				
	EST.				GI:31069921

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Homo sapiens (human)						
Homo sapiens						
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
1 (bases 1 to 958)						
L., W. B., Gruber, C., Jesses, J. and Polyes, D.						
Full-length cDNA libraries and normalization						
Unpublished						
On Feb 13, 2001 this sequence version replaced gi:12795582.						

cg1-bjhn@cluster.cgi1.peeq-CS0DM004DD01NP1&cluster=6809.f. Contact :
Feng Liang Email: liang01@feteach.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DM004DD01NP1.

FEATURES	Location/Qualifiers
source	1. .958

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004YH02"
/tissue="FETAL LIVER"
/dev_stage="fetal"
/clone_1lb="Homo sapiens FETAL LIVER"
/note="Organ: Liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end was enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT      229 a      272 g      210 t      7 others
240 a

```

BASE COUNT	240 a.	229 c	272 g	210 t	7 others
ORIGIN					
Query Match	39.9%	Score 906.2;	DB 9;	Length 958;	
Best Local Similarity	98.1%;	Pred. No. 4.8e-13;			
Matches 942;	Conservative 4;	Mismatches 11;	Indels 3;	Gaps	

QY	124	ATTGTAATTSTGGAAAAGACAGCGTGGCCAGGCTCAATTCAGAGAGGCCCAACCTTGAGCAAG	1
Db	958	AATTGTAATTSTGGAAAAGACAGCGTGGCCAGGCTCAATTCAGAGAGGCCCAACCTTGAGCAAG	8
QY	1305	CTGCGCGCCATTACTGTGCTCTCCAGCCCTTCTACTATTGTGTGCAGACAGACATCACTG	13
Db	898	CTGCGCGCCATTACTGTGCTCTCCAGCCCTTCTACTATTGTGTGCAGACAGACATCACTG	8
QY	1365	CTCTTCATGGGTTAATCTGCATGACTGCTTCTGCTCTTCACTGGGACCAATATGGCTTAA	14
Db	839	CTCTTCATGGG-TAATCTGCATGACTGCTTCTGCTCTTCACTGGGACCAATATGGCTTAA	78
QY	1425	GTGTATTAATTCATCTATTCTTGTGGCACAATCTTCTCTGAGCTACTACTATTATATG	14
Db	780	GTGTATTAATTCATCTATTCTTGTGGCACAATCTTCTCTGAGCTACTACTATTATATG	72
QY	1485	CCTTAATATTCACAAAGCAATGTGTCCAGAGAAAGAGATTAAAGAGATGAAATATCC	15
Db	720	CCTTAATATTCACAAAGCAATGTGTCCAGAGAAAGAGATTAAAGAGATGAAATATCC	66
QY	1545	ATTTCCTGTGTGCTGTGCGGGACGTGTGCACAATCTACTGTCCTCCCTTTCAAGCA	16
Db	660	ATTTCCTGTGTGCTGTGCGGGACGTGTGTGCABAATCTACTGTCCTTTTCAAGCA	66
QY	1605	CTCCTTTGCCCCAGACAGAGATGAAAAAGCCAGGAGGTGAGAGATCGATGCTTCCAG	16
Db	600	CTCCTTTGCCCCAGACAGAGATGAAAAAGCCAGGAGGTGAGAGATCGATGCTTCCAG	5
QY	1665	CTGTGCTCTGTGTGCCAGCCAAATCTTCAATTGTGGGCAAAAGGGGAAAC-TTTTTTTTGG	17
Db	540	CTGTGCTCTGTGTGCCAGCCAAATCTTCAATTGTGGGCAAAAGGGGAAAC-TTTTTTTTGG	46

QY 1724 AGAAGCGCTGCTTGTGTCACCGCAGTGAATGAGTGGGGATCTCAGCTACCCG 1783
 DB 480 AGAAGCGCTGCTTGTGTCACCGCAGTGAATGAGTGGGGATCTCAGCTACCCG 421
 QY 1784 AACCTCACCCTCTGGGTTCAAGTGAATTTCTGCTGACCTCCCAAGTACGTGGAA 1843
 DB 420 AACCTCACCCTCTGGGTTCAAGTGAATTTCTGCTGACCTCCCAAGTACGTGGAA 361
 QY 1844 ACAGGACCGCCACCATGACCCAGCTAATTTTGTATTTTGAAGTGAAGCGGATTTGACCA 1903
 DB 360 ACAGGACCGCCACCATGACCCAGCTAATTTTGTATTTTGAAGTGAAGCGGATTTGACCA 301
 QY 1904 CGTTGGCCAGCGTGTCTGGAATCTCTGACCCGCAAGTATCCACCCGCTCGGCTCCCA 1963
 DB 300 CGTTGGCCAGCGTGTCTGGAACHCTGACCCGCAAGTATCCACCCGCTCGGCTCCCA 241
 QY 1964 AAGTGTGGGATTACAGGGGTGAGCAAGTGTCCGCGCCCAAGGGAATCTGTGGG 2023
 DB 240 AAGTGTGGGATTACAGGGGTGAGCAAGTGTCCGCGCCCAAGGGAATCTGTGGG 181
 QY 2024 AGAGCAGAGGGGCTCAGATCTCCCTGTGATTCGCCATGACATGCTTATCTCTCC 2083
 DB 180 AGAGCAGAGGGGCTCAGATCTCCCTGTGATTCGCCATGACATGCTTATCTCTCC 121
 QY 2084 CCATCTAGCCAGGAATCTATGTTGTTTCTTCTGCAATTAATGATGATGATG 2143
 DB 120 CCATCTAGCCAGGAATCTATGTTGTTTCTTCTGCAATTAATGATGATGATG 61
 QY 2144 CGGCTACCAACACCCCGCATGGGGGGTGAAGAGGGGCTGCTGCTCA 2203
 DB 60 CGGCTACCAACACCCCGCATGGGGGGTGAAGAGGGGCTGCTGCTCA 1

RESULT 8
 AL557774 1201 bp mRNA linear EST 31-MAY-2003
 LOCUS AL557774 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0D003YA18 5-PRIME, mRNA sequence.
 ACCESSION AL557774
 VERSION AL557774.2 GI:31279574
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 15, 2001 this sequence version replaced gi:12901710.
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6809.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi:bin/cluster.cgi?seq=CS0D003BA09P1&cluster=6809.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D003BA09P1.
 FEATURES
 source Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D003YA18"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_1b="Homo sapiens T CELLS (JURKAT CELL LINE) COT
 10-NORMALIZED"

QY 1724 AGAAGCGCTGCTTGTGTCACCGCAGTGAATGAGTGGGGATCTCAGCTACCCG 1783
 DB 480 AGAAGCGCTGCTTGTGTCACCGCAGTGAATGAGTGGGGATCTCAGCTACCCG 421
 QY 1784 AACCTCACCCTCTGGGTTCAAGTGAATTTCTGCTGACCTCCCAAGTACGTGGAA 1843
 DB 420 AACCTCACCCTCTGGGTTCAAGTGAATTTCTGCTGACCTCCCAAGTACGTGGAA 361
 QY 1844 ACAGGACCGCCACCATGACCCAGCTAATTTTGTATTTTGAAGTGAAGCGGATTTGACCA 1903
 DB 360 ACAGGACCGCCACCATGACCCAGCTAATTTTGTATTTTGAAGTGAAGCGGATTTGACCA 301
 QY 1904 CGTTGGCCAGCGTGTCTGGAATCTCTGACCCGCAAGTATCCACCCGCTCGGCTCCCA 1963
 DB 300 CGTTGGCCAGCGTGTCTGGAACHCTGACCCGCAAGTATCCACCCGCTCGGCTCCCA 241
 QY 1964 AAGTGTGGGATTACAGGGGTGAGCAAGTGTCCGCGCCCAAGGGAATCTGTGGG 2023
 DB 240 AAGTGTGGGATTACAGGGGTGAGCAAGTGTCCGCGCCCAAGGGAATCTGTGGG 181
 QY 2024 AGAGCAGAGGGGCTCAGATCTCCCTGTGATTCGCCATGACATGCTTATCTCTCC 2083
 DB 180 AGAGCAGAGGGGCTCAGATCTCCCTGTGATTCGCCATGACATGCTTATCTCTCC 121
 QY 2084 CCATCTAGCCAGGAATCTATGTTGTTTCTTCTGCAATTAATGATGATGATG 2143
 DB 120 CCATCTAGCCAGGAATCTATGTTGTTTCTTCTGCAATTAATGATGATGATG 61
 QY 2144 CGGCTACCAACACCCCGCATGGGGGGTGAAGAGGGGCTGCTGCTCA 2203
 DB 60 CGGCTACCAACACCCCGCATGGGGGGTGAAGAGGGGCTGCTGCTCA 1

/note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 229 a 315 c 282 g 304 t 71 others

Query Match 38.4%; Score 873.4; DB 9; Length 1201;
 Best Local Similarity 91.0%; Pred. No. 2.2e-12;
 Matches 956; Conservative 32; Mismatches 50; Indels 12; Gaps 6;

QY 29 TTGGGGGGGGGGGCTGGGGGGGGTCCCTGTTGGGGGCTCCGGAGTTAAGATGAGCGGCTC 88
 DB 13 TCCCGGGAATCCGGAAGTGGGGGGTCCCTGTTGGGGGCTCCGGAGTTAAGATGAGCGGCTC 72
 QY 89 ACCGGAGGGGAGAGAGGGGAGTGTGTGGCGCTGGCGGGGGTTCGACGTGGGTTTCCA 148
 DB 73 ACCGGAGGGGAGAGAGGGGAGTGTGTGGCGCTGGCGGGGGTTCGACGTGGGTTTCCA 132
 QY 149 GGAGCTGAGCTTTAACAAGTTGGCGAGTCCCTGGGGCGGCTCAGAACAGGGCGCTGGGCT 208
 DB 133 GGAGCTGAGCTTTAACAAGTTGGCGAGTCCCTGGGGCGGCTCAGAACAGGGCGCTGGGCT 192
 QY 209 GATCATCTGCATCTTCCCTGGGGTTACCCCTTGTGTTTATGCGGATTAACCTTTCTTA 268
 DB 193 GATCATCTGCATCTTCCCTGGGGTTACCCCTTGTGTTTATGCGGATTAACCTTTCTTA 252
 QY 269 CAAGGAGACCTATCTATCAGCTCTTCCATACCTTTACAGGCTCTCAATGCTTAAT 328
 DB 253 CAAGGAGACCTATCTATCAGCTCTTCCATACCTTTACAGGCTCTCAATGCTTAAT 312
 QY 329 TAACCTTTGGAACCAAGCTCTTACCACTCCCTGCTGTGTATGTGCTTCACTTCTCATCT 388
 DB 313 TAACCTTTGGAACCAAGCTCTTACCACTCCCTGCTGTGTATGTGCTTCACTTCTCATCT 372
 QY 389 TCGACTAATGGGCGGACCATCATCTGCGGCTGCTCACTACCTTTGCTTCCAGATGGGCTTA 448
 DB 373 TCGACTAATGGGCGGACCATCATCTGCGGCTGCTCACTACCTTTGCTTCCAGATGGGCTTA 432
 QY 449 CCTTCTGGCTGATATCTATTTACACTGCCACCGGCACTACGATATCAAGTGAACAATGCC 508
 DB 433 CCTTCTGGCTGATATCTATTTACACTGCCACCGGCACTACGATATCAAGTGAACAATGCC 492
 QY 509 ACATTTGTTTCTGACTTTTGAACCTGATTTGGCTGTGACTTACTTGAAGAGGAGGA 568
 DB 493 ACATTTGTTTCTGACTTTTGAACCTGATTTGGCTGTGACTTACTTGAAGAGGAGGA 552
 QY 569 ACATCAGAAATCTTCTGCTCTGAGCAACAGAAATATGACATACGTTGGTCTTCCCT 628
 DB 553 ACATCAGAAATCTTCTGCTCTGAGCAACAGAAATATGACATACGTTGGTCTTCCCT 612
 QY 629 GCTGGAAGTTGCTGTTTCTCTACTTCTATGAGGAGCTTCTGTAAGGAGCCCAAGTTCTC 688
 DB 613 GCTGGAAGTTGCTGTTTCTCTACTTCTATGAGGAGCTTCTGTAAGGAGCCCAAGTTCTC 672
 QY 689 AATGAATCTATCATGAAGCTGTGTGAGGAGAGCTGATGACATACCGAGAAATATAC 748
 DB 673 AATGAATCTATCATGAAGCTGTGTGAGGAGAGCTGATGACATACCGAGAAATATAC 732
 QY 749 AAACAGCATCATCTCTGCTCTGAGCGCTGAGGCTGAGGCTTTTCTACTAGTGGGCTA 808
 DB 733 AAACAGCATCATCTCTGCTCTGAGCGCTGAGGCTGAGGCTTTTCTACTAGTGGGCTA 792
 QY 809 CACACTGTCAAGCCCGCACATGACAGAGACTATCTCTCACTGAAGACTATGACACCA 868
 DB 793 CACACTGTCAAGCCCGCACATGACAGAGACTATCTCTCACTGAAGACTATGACACCA 851
 QY 869 CCCCTTGTGTTCCGCTGATATCATGCTATCTGGGGCAAGTTGTGCTGTACAAATA 928
 DB 852 CCCCTTGTGTTCCGCTGATATCATGCTATCTGGGGCAAGTTGTGCTGTACAAATA 907
 QY 929 TGTCACTGTGCTGTGACAGAGAGATGATTTTGAAGGAGCTGGGCTTCAATAG 988

Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6809.f. For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ003BA09NP1&cluster=6809.f. Contact :
 Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DJ003BA09NP1.

FEATURES

source
 1.1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ003YAL8"
 /cell_type="T CELLS (URKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="URKAT"
 /clone_1ib="Homo sapiens T CELLS (URKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 258 a 247 c 334 g 250 t 112 others
 ORIGIN

Query Match 37.7%; Score 856.2; DB 9; Length 1201;
 Best Local Similarity 90.2%; Pred. No. 5.2e-12;
 Matches 881; Conservative 50; Mismatches 42; Indels 4; Gaps 3;

1216 ACCGTGCTGCTTCAGATGAATTCCTATTGTTATGTGGAAGACAGCTGCCAGGC 1275
 976 MNCGGTCKGYCTCYGA--GGAATCCTCCATCTGTTGTTGAAACACDKGKCCRC 919
 1276 TCATTGAAGAGCCGACCCAGAGAGCTGCCGACATTAATGCTCCAGCCCTTCT 1335
 918 TCATCCARKRTACCCGACCCCTAGAGAGKGGCCGACATWAKTTCCTCCASCTCYT 859
 1336 ACTATTGTGTCAGACAGACCATCAGCTGCTCTTCATAGGTTACTCCATGACTGCT 1395
 858 ACTATTGTGTCAGACAGACCATCAGCTGCTCTTCATAGGTTACTCCATGACTGCT 799
 1396 GCTCTTTCAGCTGGAACAAATGCTTAAGGTATTAATCATCTATTCTTCTGG--CCAC 1454
 798 GCCTCTTTCAGCTGGAACAAATGCTTAAGGTATTAATCATCTATTCTTCTGGSCAC 739
 1455 ATCTCTCTGCGCTTCTATTCATATTCCTTATTCACAAAGCATGTCGCAAGG 1514
 738 ATCTCTCTGCGCTTCTATTCATATTCCTTATTCACAAAGCATGTCGCAAGG 679
 1515 AAAGAGAGTTAAAGAGATGATATATCCATTTCCTGCTGCTGCGGAGCTGCTG 1574
 678 AAAGAGAGTTAAAGAGATGATATATCCATTTCCTGCTGCTGCGGAGCTGCTG 619
 1575 CAGAACTACTGCTGCTCTTTCACAGCACTCTTTCGCCGAGAGAGAAATGAAAA 1634
 618 CAGAACTACTGCTGCTCTTTCACAGCACTCTTTCGCCGAGAGAGAAATGAAAA 559
 1635 GCCAGGAGGTGAGATGATGATCTTCAGCTGCTGCTGCTGCGCAGGCAAGCTTCT 1694
 558 GCCAGGAGGTGAGATGATGATCTTCAGCTGCTGCTGCTGCGCAGGCAAGCANNCA 499
 1695 TTGGGCGCAAAAGGGAATTTTGTGAGAGAGGCTCTGCTTGTGCAACCCAGCTGG 1754
 498 TTGGGCGCAAAAGGGAATTTTGTGAGAGAGGCTCTGCTTGTGMAACCAAGCTGG 439
 1755 AATGCACTGCGGAGATCTACCTGACCGCACTCTGCTGCTGCTGCTGCTGCTGCT 1814
 438 AATGCACTGCGGAGATCTACCTGACCGCACTCTGCTGCTGCTGCTGCTGCTGCTG 379
 1815 CTGCTCAGCTCTCCAAAGTATGCTGGAATACAGGACGCAACATGCCCAATTTT 1874

DB 378 CTGCTCAGCTCTCCAAAGTATGCTGGAAMACGGAACGCCACACATGCCAGCAATTTT 319
 1875 GTATTTTTCAGTAAAGGAGATTTTACACAGTGTGCGCAGGCTGTCTTCACTCTGACC 1934
 318 GTATTTTTCAGTAAAGGAGATTTTACACAGTGTGCGCAGGCTGTCTTCACTCTGACC 259
 1935 GCAAGTATCCACCCGCTCTCCCTCCCAAGTGTGGATTAACAGGCTGACCCAGCT 1994
 258 GCAAGTATCCACCCGCTCTCCCTCCCAAGTGTGGATTAACAGGCTGACCCAGCT 199
 1995 GCCCGGCCCCAAAGGGAATCTTGGGAGAGAGAGAGGAGCTACATCTCCCTCTGA 2054
 198 GCCCGGCCCCAAAGGGAATCTTGGGAGAGAGAGAGGAGATCACTCTCCCTCMA 139
 2055 TTCCCGCATGACATTTGCTTCTCTCCCATCTAGCCAGGATCTATTGTTTCT 2114
 138 TTCCCGCATGACATTTGCTTCTCTCCCGMAAMAGCAGGAATCTATTGTTTCT 79
 2115 TCTGCCAATTACTATGATTTGTTATGTCGCTTACACACACCCCTCATGGGGGCTG 2174
 78 TCTGCCAATTACTATGATTTGTTATGTCGCTTACACACACCCCTCATGGGGGCTG 20
 2175 GAGAGGGGTCGCAAGGCC 2191
 DB 19 GAGAGGGGTCGCAAGGCC 3

RESULT 11
 AL576170/c 1201 bp mRNA linear EST 01-JUN-2003
 LOCUS AL576170 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0DI072Y019 3-PRIME, mRNA sequence.
 ACCESSION AL576170.2 GI:31314460
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Li,W.B., Gruber,C., Jessee,J. and Polayer,D.
 Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 16, 2001 this sequence version replaced gi:12938048.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6809.f. For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DI072AH10NP1&cluster=6809.f. Contact :
 Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DI072AH10NP1.

FEATURES

source

1.1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DI072Y019"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT

284 a 261 c 322 g 266 t 68 others
 ORIGIN
 Query Match 37.6%; Score 855.2; DB 9; Length 1201;

Best Local Similarity 95.0%; Pred. No. 5.5e-12;
Matches 896; Conservative 13; Mismatches 30; Indels 4; Gaps 3;

```

QY 1303 AGCTGGCCGCACTTACTGCTCCCTCCAGCCCTTCTACTATTGGTGCACAGACCATTCCTACT 1362
DB 943 AGMAYSTGGCCSCATPAATGCTCCCTCCAGCCCTTCTACTATTGGTGCACAGACCATTCCTACT 884
QY 1363 GGCCTTCATGAGGTTACTCATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422
DB 883 GGCCTTCATGAGGTTACTCATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824
QY 1423 AGGTGATPAATTCATCTATTTT-CCTGGCCACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1481
DB 823 AGGTGATPAATTCATCTATTTTCCCTGGCCACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
QY 1482 TTGCTCTTATTCACAAAGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1541
DB 763 TTGCTCTTATTCACAAAGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
QY 1542 TCCATTTCCCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1601
DB 703 TCCATTTCCCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
QY 1602 GCATCTCTTTGCCCCAGAGAGAGATGAGAAAGCCAGGAGGTGAGATGATGCTTTC 1661
DB 643 GCATCTCTTTGCCCCAGAGAGAGATGAGAAAGCCAGGAGGTGAGATGATGATGCTTTC 584
QY 1662 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1721
DB 583 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 524
QY 1722 GGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1781
DB 523 GGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464
QY 1782 GCAACCTCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1841
DB 463 GCAACCTCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
QY 1842 ATACAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1901
DB 403 ATACAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
QY 1902 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1961
DB 343 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 284
QY 1962 CAAAGTGTGAGATTACAGGCTGAGCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2021
DB 283 CAAAGTGTGAGATTACAGGCTGAGCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224
QY 2022 GGAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2081
DB 223 GGAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164
QY 2082 CCCCATCTAGCAGGAATCTATTTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2141
DB 163 CCCCATCTAGCAGGAATCTATTTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104
QY 2142 TGGCGCTACACCA--CCCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCTGCTGCTGCTGCTGCTGCT 2199
DB 103 TGGCGCTACACCAAGGCGCCGCCCATGGGGGGGTGAGAGGGGTGCAAGGCTGCTGCTGCTGCTGCTGCT 44
QY 2200 TCCAC-TTTTCTTACTCTTGAAGCTGATTAATAATCACTT 2241
DB 43 TCCAC-TTTTCTTACTCTTGAAGCTGATTAATAATCACTT 1

```

RESULT 12
BMS53135 1097 bp mRNA 1linear EST 20-FEB-2002
LOCUS BMS53135
DEFINITION AGENCOURT_6572558 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467062
5', mRNA sequence.

```

ACCESSION BMS53135
VERSION BMS53135.1 GI:18791600
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 1097)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM1968 row: h column: 07
High quality sequence stop: 666.
Location/Qualifiers
1..1097
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5467062"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_41"
/note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 244 a 315 c 254 g 275 t 9 others
ORIGIN
Query Match 37.1%; Score 843.2; DB 12; Length 1097;
Best Local Similarity 94.5%; Pred. No. 1.1e-11;
Matches 937; Conservative 0; Mismatches 41; Indels 14; Gaps 6;
QY 1092 ATCAACACCAACGCTGGGTGCGCCGCTACATCTTCAACGACTCAAGTTCCTTGGAAAT 1151
DB 1 ATCAACACCAACGCTGGGTGCGCCGCTACATCTTCAACGACTCAAGTTCCTTGGAAAT 60
QY 1152 AAGAACTCTCTCAGGGTCTCTGTTGCTATTCTGCGCCCTTGGCAGGCTGCACTCA 1211
DB 61 AAGAACTCTCTCAGGGTCTCTGTTGCTATTCTGCGCCCTTGGCAGGCTGCACTCA 120
QY 1212 GGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
DB 121 GGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 1272 AGGCTCATTCAGAGAGCCCACTCTGAGCAAGCTGCGCGCCATTAAGTCTCTCAGGCC 1331
DB 181 AGGCTCATTCAGAGAGCCCACTCTGAGCAAGCTGCGCGCCATTAAGTCTCTCAGGCC 240
QY 1332 TTCTACTATTGGTGCACAGACCATTCACATGCTCTTCAATGAGTTACTCCATGACTGCC 1391
DB 241 TTCTACTATTGGTGCACAGACCATTCACATGCTCTTCAATGAGTTACTCCATGACTGCC 300
QY 1392 TTCTGCTCTTCAAGTGGAGCAATGAGTAAAGTAAATCCATCTATTCTTCTGGC 1451
DB 301 TTCTGCTCTTCAAGTGGAGCAATGAGTAAAGTAAATCCATCTATTCTTCTGGC 360
QY 1452 CACATCTTCTCTGAGGCTACTATTCATATTTGCTTATTTACAAAGCAATGCTGCCA 1511
DB 361 CACATCTTCTCTGAGGCTACTATTCATATTTGCTTATTTACAAAGCAATGCTGCCA 420

```

QY	1512	AGGAAAGAGAGAGTTAAAMAAATGGAATAATTCATTCTCCCTGGATGGCTGTGCGGGACTG	1577
Db	421	AGGAAAGAGAGAGTTAAAGAAATGGAATAATTCATTCTCCCTGGATGGCTGTGCGGGACTG	480
QY	1572	GTGCAAGAAACTACTGTCGTCTCCCTTTTCCACAGCACTCCTTTGGCCGACAGAGAAATGGA	1631
Db	481	GTGCAAGAAACTACTGTCGTCTCCCTTTTCCACAGCACTCCTTTGGCCGACAGAGAAATGGA	540
QY	1632	AAACCCAGGGAGGTGGAGATGATGATGCTTCCAGCTGTGTCCTTGTGTGCGACGCAAGTCTT	1691
Db	541	AAACCCAGGGAGGTGGAGATGATGATGCTTCCAGCTGTGTCCTTGTGTGCGACGCAAGTCTT	600
QY	1692	CATTGGGGGCGAAAGGGGAACTTTTTTTTGGAGAAAGCGTCTTGCTTTGTCAACCCAGC	1751
Db	601	CATTGGGGGCGAAAGGGGAACTTTTTTTTGGAGAAAGCGTCTTGCTTTGTCAACCCAGC	660
QY	1752	TGGAATGCAGTGGCGGGATCTCAGCTCACCGCAACCTCCACTCCTGGGTTCAATGATTT	1811
Db	661	TGGAATGCAGTGGCGGGATCTCAGCTCACCGCAACCTCCACTCCTGGGTTCAATGATTT	720
QY	1812	TTCTGTGCTTCAAGCTTCCCAAGTAGCTGGGAATACAGGCAAGCCACATGCCCCAGTAAAT	1871
Db	721	TTCTGTGCTTCAAGCTTCCCAAGTAGCTGGGAATACAGGCAAGCCACATGCCCCAGTAAAT	779
QY	1872	TTTGATTTTTCAGTAGAAGACGGGATTTTACACAGTCGTGGCCAGGC-TGTCTTCGAATCTCT	1933
Db	780	TTTGATTTTTCAGTAGAAGACGGGATTTTACACAGTCGTGGCCAGGC-TGTCTTCGAATCTCT	839
QY	1931	GACCGCAAGTAGATCACCC--GCTCTCCGCTCTCCAAATGCT-GGGATTACAGGCGT--G	1985
Db	840	GACCGCAAGTAGATCCANCCGNCCTCCGNNCTCCCAAATGCTGGGGATTACAGGCGGATGA	899
QY	1986	AGCCACCGTGCCTCCGGCCCAAGAGGGAATCTTGTG-----GGAGGACGAGAGGGCT	2038
Db	900	AGCCACCGTGCCTCCGGCCCAAGAGGGAATCTTGTGTGGAAGAGCAAGAGGGGCTCA	959
QY	2039	CACATCTCCCTCTGATTTCCCCCATGCAATT 2070	
Db	960	CATCTTCCCTCTGTGATTTCCCCCATGCAATT 991	

RESULT, 13	LOCUS	DEFINITION	EST 15-MAY-2003
BX431852	BX431852	876 bp	mRNA
BX431852	BX431852	Homo sapiens FETAL LIVER Homo sapiens	linear
CSD0M004YH02.5-PRIME.	CSD0M004YH02.5-PRIME.	mRNA sequence.	cdna clone

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 876)
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Playes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
COMMENT	Contact: Genoscope

COMMENT

Contact: Genoscope
Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6809.f. For
more information about this cluster, see

cgj-bin/cluster.cgi?seq=CS0BAAG0655Z03_CS06165_1&cluster=6809.f
Contact : Peng Liang Email : liang@life.uchicago.edu
http://fulllength.invitrogen.com/Invitrogen/Corporation_1600
Faraday Avenue Genoscope Sequence ID : CS0BAAG0655Z03_CS06165_1

```

FEATURES
source      Location/Qualifiers
1.  .876    /organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM004YH02"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT      207 a      237 c      201 g      227 t      4 others
ORIGIN

```

Query Match	36.8%	Score 837.6;	DB 13;	Length 876;
Best Local Similarity	99.0%	Pred. No. 1.7e-11;		
Matches 862;	Conservative 0;	Mismatches 7;	Indels 2;	Gaps 2

QY	842	TTCTCTCAGTGAAGACATATAGACAACACCCTTGTGGTTCCGCTGCAGATGTACATGCTGAT	901
Db	3	TTCTCTCAGTGTGAAGCCTATGACAAACACCCTTGTGGTTCCGCTGCAGATGTACATGCTGAT	62
QY	902	CTGGGCGCAAGTTGTGTGCTGTACAAATATATGTACCTGTGGCTGTGTCACAGAAAGATATG	961
Db	63	CTGGGCGCAAGTTGTGTGCTGTACAAATATATGTACCTGTGGCTGTGTCACAGAAAGATATG	122
QY	962	CATTTTAACGGGCGCTGGGGCTTTCATATGGCTTTGAAAGAAAGGGCAAGCAAGTGGGATGC	1021
Db	123	CATTTTAACGGGCGCTGGGGCTTTCATATGGCTTTGAAAGAAAGGGCAAGCAAGTGGGATGC	182
QY	1022	CTGTGCCAAGATGAAAGGTGTGGCTCTTTTGAAGCAAAACCCCGCTTCACTGGCACCATTGC	1081
Db	183	CTGTGCCAAGATGAAAGGTGTGGCTCTTTTGAAGCAAAACCCCGCTTCACTGGCACCATTGC	242
QY	1082	CTCATTTCAACATCAACACCAAACCGCTGGGTGGCCCGCTCATCTTTCAAGCACTCAAGTT	1141
Db	243	CTCATTTCAACATCAACACCAAACCGCTGGGTGGCCCGCTCATCTTTCAAGCACTCAAGTT	302
QY	1142	CCTTGGAATTAAGAACTCTCTCAGGGGCTCTGTTGTAATTCCTGGCCCTCTGGACAGG	1201
Db	303	CCTTGGAATTAAGAACTCTCTCAGGGGCTCTGTTGTAATTCCTGGCCCTCTGGACAGG	362
QY	1202	CCTGCACCTCAGAGATACCTGTCTGTGCTTCCAGATGAAATTCCTCATTTGTATTGTGGAAAG	1261
Db	363	CCTGCACCTCAGAGATACCTGTGTCTGCTCTTCCAGATGAAATTCCTCATTTGTATTGTGGAAAG	422
QY	1262	ACAGGGCTGCCAGGCTCATTTCAAAGAGACCCCACTTGAGCAAGCTGGCCCGCATTACTGT	1321
Db	423	ACAGGGCTGCCAGGCTCATTTCAAAGAGACCCCACTTGAGCAAGCTGGCCCGCATTACTGT	482
QY	1322	CCTGCACGCCCTTCTACTATTGTTGGTGAACAAGACCATGCACTGGGCTCTTCATAGGGTTACTC	1381
Db	483	CCTGCACGCCCTTCTACTATTGTTGGTGAACAAGACCATGCACTGGGCTCTTCATAGGGTTACTC	542
QY	1382	CATGACTGCCTTCTGTCCCTCTTCACGTGGGACAAATGSGCTTAAGGTGTATAAATCCATCTA	1441
Db	543	CATGACTGCCTTCTGTCCCTCTTCACGTGGGACAAATGSGCTTAAGGTGTATAAATCCATCTA	602
QY	1442	TTTTCCCTTGGCCACATCTTTTCCCTGAGACCTATATTCATATGGCTTATATTTCAACAAGC	1501
Db	603	TTTTCCCTTGGCCACATCTTTTCCCTGAGACCTATATTCATATATGGCTTATATTTCAACAAGC	662
QY	1502	AATGATGCCAAGAAAGAAAGAAAGTAAAGAAAGATGGAATTAATCCATTTTCCCTGTGGGCTGT	1561
Db	663	AATGATGCCAAGAAAGAAAGAAAGTAAAGAAAGATGGAATTAATCCATTTTCCCTGTGGGCTGT	722
QY	1562	TGCGGGAATGTGTGCAAGAACTACTGTCTCTCCCTTTTCAACAGCACTCCTTGGCCCCAGAGC	1621
Db	723	TGCGGGAATGTGTGCAAGAACTACTGTCTCTCCCTTTTCAACAGCACTCCTTGGCCCCAGAGC	782
QY	1622	AGAAATATGAAATAAACCCAGGAGAGGT-GGAAATATGATGCTTCCACTGTGGCCTCTGTGCC	1681
Db	783	AGAAATATGAAATAAACCCAGGAGAGGTGAGGAAGTTCATGCTTTNACACTGTGCCCTCTGTGCC	842

QY 1681 ACCCAAGTCTTATTTGGGGCCAAAGGGAA 1711
| | | | |
Db 843 AGCCAAGTC-TCAATTTGGGCCAAAGGAAA 872

RESULT 14
BQ917856 974 bp mRNA linear EST 20-AUG-2002
LOCUS BQ917856
DEFINITION AGENCOURT 8801447 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6311255
5' mRNA sequence.
ACCESSION BQ917856
VERSION BQ917856.1 GI:22332554
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Makiyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL 1 (bases 1 to 974)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
cDNA Library Preparation: Resgen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM13734 row: n column: 24
High quality sequence stop: 755.
Location/Qualifiers
1..974
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6311255"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_129"
/note="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1.cdb; Site_1: EcoRV; Site_2: NotI; Cloned
unidirectionally. Primer: Oligo dt. Average insert size
2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this
is a NIH MGC library."

BASE COUNT 199 a 260 c 248 g 267 t
ORIGIN

Query Match 36.8%; Score 836.4; DB 13; Length 974;
Best Local Similarity 97.2%; Pred. No. 1.7e-11;
Matches 883; Conservative 0; Mismatches 21; Indels 4; Gaps 3;

QY 33 GGGCGCGGAGCTGGGGGGGCTCTGTGGGCTCCCGAGTTAAGATGGCTCTCAGCG 92
| | | | |
Db 1 GGTACCGAGGCTGTGGTGGGAGATTGCGACAGAGGAGATTAGATGGCGTCTCAGCG 60

QY 93 GAGGGGAGCAGAGGAGCTGTGGGCGCGGGGGGTTCTCAGTCGGGTTTCCAGAG 152
| | | | |
Db 61 GAGGGGAGCAGAGGAGCTGTGGGCGCGGGGGTCTCAGTCGGGTTTCCAGAG 120

QY 153 CTGAGCCTTAAACAATTGGCGAGTCCCTGGGCGCGTCAAGACAGGCGCTGGCGCTGATC 212
| | | | |
Db 121 CTGAGCCTTAAACAATTGGCGAGTCCCTGGGCGCGTCAAGACAGGCGCTGGCGCTGATC 180

QY 213 ATCTTCATCTTCTGGGTTACCCCTTGTGTTTATCGCATTTACTTTTTCACAG 272
| | | | |
Db 181 ATCTTCATCTTCTGGGTTACCCCTTGTGTTTATCGCATTTACTTTTTCACAG 240

QY 273 GAGACTACCTCATCACTCTTCACTTACCTTACAGGCGCTCAATGCTATTATTAAC 332
| | | | |
Db 241 GAGACTACCTCATCACTCTTCACTTACCTTACAGGCGCTCAATGCTATTATTAAC 300

QY 333 TTGGAAACACAGCTCTACCACTCCCTGCTGTGTATTTGCTCACTTCTCATCTTGA 392
| | | | |
Db 301 TTGGAAACACAGCTCTACCACTCCCTGCTGTGTATTTGCTCACTTCTCATCTTGA 360

QY 393 CTATGGCCCGACCATCATCTGCGCTCTCACTACCTTTTGGCTTCCAGATGGCTTACTT 452
| | | | |
Db 361 CTATGGCCCGACCATCATCTGCGCTCTCACTACCTTTTGGCTTCCAGATGGCTTACTT 420

QY 453 CTGGGTGATCTATTACCTGCGACCGGCAACTACGATACATCAATGACAAATGACAT 512
| | | | |
Db 421 CTGGGTGATCTATTACCTGCGACCGGCAACTACGATACATCAATGACAAATGACAT 480

QY 513 TGTGTTGACTTTGAACTGATTGTTGGCTGTGACTACTTTGACGAGGAGAAAGAT 572
| | | | |
Db 481 TGTGTTGACTTTGAACTGATTGTTGGCTGTGACTACTTTGACGAGGAGAAAGAT 540

QY 573 CAGAAATTCCTGCTCTGAGCAACAGAAATATGCAATACGTGTTCTTCCCTGCTG 632
| | | | |
Db 541 CAGAAATTCCTGCTCTGAGCAACAGAAATATGCAATACGTGTTCTTCCCTGCTG 600

QY 633 GAACTTGTGCTGTTCTCTACTCTTATGAGGCGCTTGTGTAAGGCGCCAGTTCTCAATG 692
| | | | |
Db 601 GAACTTGTGCTGTTCTCTACTCTTATGAGGCGCTTGTGTAAGGCGCCAGTTCTCAATG 660

QY 693 AATCACTACATGAAGCTGTGACAGGAGAGCTGATTGACATACAGAAAGATACCAAC 752
| | | | |
Db 661 AATCACTACATGAAGCTGTGACAGGAGAGCTGATTGACATACAGAAAGATACCAAC 720

QY 753 AGCATATTCCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTTACCTGATGGGCTACACA 812
| | | | |
Db 721 AGCATATTCCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTTACCTGATGGGCTACACA 780

QY 813 CTGCTCACCCCGCATACAGAGATATCTCTCATGAAAGATATAGACA-CCACC 871
| | | | |
Db 781 CTGCTCACCCCGCATACAGAGATATCTCTCATGAAAGATATAGACA-CCACC 840

QY 872 CTTCGTGTTCCGCTGATGATCAT-GCTGATCTGGGG--CAAGTTGTGCTTACAAATA 928
| | | | |
Db 841 CTTCGTGTTCCGCTGATGATCATGCTGATCTGGGGCCAAAGTTGTGCTTACAAATA 900

QY 929 TGTCACT 936
| | | | |
Db 901 TGGCACCT 908

RESULT 15
B1767794 882 bp mRNA linear EST 25-SEP-2001
LOCUS B1767794
DEFINITION 603060787F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5210300 5',
mRNA sequence.
ACCESSION B1767794
VERSION B1767794.1 GI:15759372
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Makiyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL 1 (bases 1 to 882)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11527 row: m column: 21
High quality sequence stop: 840.
Location/Qualifiers

source 1. .882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5210300"
/lab_host="DH10B"
/clone_1ib="NIH MGC 122"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source:
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH MGC Library."
BASE COUNT 209 a 246 c 200 g 227 t
ORIGIN

Query Match 36.4%; Score 828.4; DB 12; Length 882;
Best Local Similarity 99.0%; Pred. No. 2.7e-11;

Matches 865; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 799 TAGTGGGCTACACACTGCTCAGCCCCCACAATCAAGAGACTATCTCTCCTGAGAGACT 858
DB 1 TAGTGGGCTACACACTGCTCAGCCCCCACAATCAAGAGACTATCTCTCCTGAGAGACT 60
QY 859 ATGACACACACCCCTTCTGCTTCGCTGCAATGATCTGAGTCTGAGGCAAGTTTGTGC 918
DB 61 ATGACACACACCCCTTCTGCTTCGCTGCAATGATCTGAGTCTGAGGCAAGTTTGTGC 120
QY 919 TGTACAAATATGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978
DB 121 TGTACAAATATGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 979 GCTTCATGCTTGTGAAGAAAGGCAAGGCAAGGATGCTGCTGCTGCTGCTGCTGCTG 1038
DB 181 GCTTCATGCTTGTGAAGAAAGGCAAGGCAAGGATGCTGCTGCTGCTGCTGCTGCTG 240
QY 1039 TGTGCTCTTTGAAACAAACCCCGCTTCACTGGCAACATTTGCTCATTCATCAACA 1098
DB 241 TGTGCTCTTTGAAACAAACCCCGCTTCACTGGCAACATTTGCTCATTCATCAACA 300
QY 1099 CCAAGCCCTGGGCGGCGGCTGATATCTTCAAGACTCAAGTTCTTTGAAATTAAGAAC 1158
DB 301 CCAAGCCCTGGGCGGCGGCTGATATCTTCAAGACTCAAGTTCTTTGAAATTAAGAAC 360
QY 1159 TCTCTCAGGCTCTCTGCTGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
DB 361 TCTCTCAGGCTCTCTGCTGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 1219 TGTGCTCTTTGCAATGAAATTCCTCATTTGTTATTTGAAAGACAGGCTGCCAGCTCA 1278
DB 421 TGTGCTCTTTGCAATGAAATTCCTCATTTGTTATTTGAAAGACAGGCTGCCAGCTCA 480
QY 1279 TTCAAGAGAGCCCGACCTGAGAAAGCTGCGGCGCAATTAATCTCTCCAGCCCTTCACT 1338
DB 481 TTCAAGAGAGCCCGACCTGAGAAAGCTGCGGCGCAATTAATCTCTCCAGCCCTTCACT 540
QY 1339 ATTGGTGAACAGACCATCACTGCTCTTCAATGGTTACTCATGACTGCTTCTGCC 1398
DB 541 ATTGGTGAACAGACCATCACTGCTCTTCAATGGTTACTCATGACTGCTTCTGCC 600
QY 1399 TCTTCAGTGGGCAAAATGCTTAAGGTATTAATCAATCAATTTCTT-TGGCCACATC 1457
DB 601 TCTTCAGTGGGCAAAATGCTTAAGGTATTAATCAATCAATTTCTTCTGCGCACATC 660
QY 1458 TTCTTCTAGCTACTATTCATATGCTTATATTCACAAAGCAATGGTCCCAAGAAA 1517
DB 661 TTCTTCTAGCTACTATTCATATGCTTATATTCACAAAGCAATGGTCCCAAGAAA 720
QY 1518 GAGAACTTAAGAAAGATGAATATCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1577

DB 721 GAGAACTT-ACGAAGATGAATATATCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 779
QY 1578 AAAGTACTGCTCTCCCTTTTACA-GCACTCTTTGCCCCAGAGCAGAGATGAAAAAGC 1636
DB 780 AAAGTACTGCTCTCCCTTTTACAAGGCACTCTTTGGCCAGAGCAGAGATGAAAAAGC 839
QY 1637 CAGGAGGTGGAAGATGATGCTTCCAGCTGTGC 1670
DB 840 CAGGAGGTGGAAGATGATGCTTCCAGCTGTGC 873

Search completed: November 22, 2003, 03:38:01
Job time : 4752 secs

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1040AB030P1.
Location/Qualifiers

FEATURES
source

1. 1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1040Y05"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 240 a 292 c 294 g 299 t 5 others

ORIGIN

Query Match 42.3%; Score 961; DB 9; Length 1130;
Best Local Similarity 100.0%; Pred. No. 1.5e-155; Indels 0; Gaps 0;
Matches 961; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GGGGGGTCCTGTTGGGCTCCGAGTTAAGTGGCTCCTCAGCGAGGGGAGCGAGG 106
DB 124 GGGGGGTCCTGTTGGGCTCCGAGTTAAGTGGCTCCTCAGCGAGGGGAGCGAGG 183
QY 107 GACTGTGTGGCCCTGCGCGGGGTTCTGCACTGGGTTTCCAGAGAGCTGAGCTTAACA 166
DB 184 GACTGTGTGGCCCTGCGCGGGGTTCTGCACTGGGTTTCCAGAGAGCTGAGCTTAACA 243
QY 167 GTTGGCGAGCTCCTGGGCGCTGAGAGAGCGGCTGATCATCTCCATCTTCT 226
DB 244 GTTGGCGAGCTCCTGGGCGCTGAGAGAGCGGCTGATCATCTCCATCTTCT 303
QY 227 GGGTTACCCCTTGTCTTTTATGCGATTAACCTTTTACAGAGAGCTTAACCTCAT 286
DB 304 GGGTTACCCCTTGTCTTTTATGCGATTAACCTTTTACAGAGAGCTTAACCTCAT 363
QY 287 CCACCTTCTTCATACCTTTACAGGCTCCTCAATGCTTATTTAATTTGAAACAGCT 346
DB 364 CCACCTTCTTCATACCTTTACAGGCTCCTCAATGCTTATTTAATTTGAAACAGCT 423
QY 347 CTACCACTCCCTGCTGTGTATGATGCTTCACTTCTCATCTTCTGACTAATGAGCGCAC 406
DB 424 CTACCACTCCCTGCTGTGTATGATGCTTCACTTCTCATCTTCTGACTAATGAGCGCAC 483
QY 407 CATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTTACCTTCTGCTGATATCTA 466
DB 484 CATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTTACCTTCTGCTGATATCTA 543
QY 467 TTACACTGCGACCGGCACTACGATATCAAGTGAAGCAATGCCATCTGTCTGACTTT 526
DB 544 TTACACTGCGACCGGCACTACGATATCAAGTGAAGCAATGCCATCTGTCTGACTTT 603
QY 527 GAAGCTGATGTTGCTGTGCTTGTGAGAGGCGCCAGTCTCAATGAATCACTACATGA 586
DB 604 GAAGCTGATGTTGCTGTGCTTGTGAGAGGCGCCAGTCTCAATGAATCACTACATGA 663
QY 587 CTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCCCTGCTGAGAGTTGCTGTT 646
DB 664 CTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCCCTGCTGAGAGTTGCTGTT 723
QY 647 CTCTGACTTTATGAGGCTTCTTGTGAGAGGCGCCAGTCTCAATGAATCACTACATGA 706
DB 724 CTCTGACTTTATGAGGCTTCTTGTGAGAGGCGCCAGTCTCAATGAATCACTACATGA 783
QY 707 GCTGTGACAGAGAGCTGATGATACATACAGAGAAATATCCAAACAGCACTACCTGTC 766
DB 784 GCTGTGACAGAGAGCTGATGATACATACAGAGAAATATCCAAACAGCACTACCTGTC 843
QY 767 TCTCAAGCGCTGAGTGTGGCTTTTCTACCTAGTGGCTACACACTGCTCAGCCCCCA 826
DB 844 TCTCAAGCGCTGAGTGTGGCTTTTCTACCTAGTGGCTACACACTGCTCAGCCCCCA 903

QY 827 CATCAAGAGACTATCTCTCACTGAGACTATGACAACCAACCCCTTGTGTTCCGCTG 886
DB 904 CATCAAGAGACTATCTCTCACTGAGACTATGACAACCAACCCCTTGTGTTCCGCTG 963
QY 887 CATGTACATGCTGATCTGGGCAAGTTTGTGCTGTAACAATATGTCACCTGTTGCTGT 946
DB 964 CATGTACATGCTGATCTGGGCAAGTTTGTGCTGTAACAATATGTCACCTGTTGCTGT 1023
QY 947 CACAGAGAGATGATGATTTTGAACGGGCTGCGCTTCAATGCTTTGAAGAAAGGCA 1006
DB 1024 CACAGAGAGATGATGATTTTGAACGGGCTGCGCTTCAATGCTTTGAAGAAAGGCA 1083
QY 1007 G 1007
DB 1084 G 1084

RESULT 2

AL532090

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL532090 1000 bp mRNA linear EST 23-MAY-2003
AL532090 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM004YH02 5-PRIME, mRNA sequence.
AL532090.2 GI:31069922
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1000)
Lj W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12795583.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6809.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM004YH01&cluster=6809.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DM004YH01P1.
Location/Qualifiers

FEATURES

source

1. 1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004YH02"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

38.8%; Score 881; DB 9; Length 1000;
Pred. No. 7.5e-142;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 GAGGGGACGAGGAGCTGTGTGGCTGCGGGGCTTCTGAGTGGGTTCCAGAG 152
DB 120 GAGGGGACGAGGAGCTGTGTGGCTGCGGGGCTTCTGAGTGGGTTCCAGAG 179
QY 153 CTGAGCTTAACAAGTTGGGAGCTCCCTGGGGCTGAGAACAGGCGCTGAGCTGATC 212

Db 618 CTCTCTCTGACGCTACTATTCATTTGCTTATTTACAAAGCAATGGTGCAGGAA 677

QY 1517 AGAAGAGTTAAAGAGATGGAATPACATTTCTCTGGTGCGCTGTGCGGAGCTGTGCA 1576

Db 678 AGAAGAGTTAAAGAGATGGAATPACATTTCTCTGGTGCGCTGTGCGGAGCTGTGCA 737

QY 1577 GAAGACTACTGCTGCTCCCTTTTTCACAGACCTCTTTGGCCCGAGAGCAGAGAAATGAAAGC 1636

Db 738 GAAGACTACTGCTGCTCCCTTTTTCACAGACCTCTTTGGCCCGAGAGCAGAGAAATGAAAGC 797

QY 1637 CAGG 1640

Db 798 CAGG 801

RESULT 4
BX339313 1149 bp mRNA linear EST 02-MAY-2003
LOCUS BX339313 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1070YJ16 5-PRIME, mRNA sequence.
ACCESSION BX339313 GI:30335787
VERSION BX339313
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6809.f For more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSOD1070DE08QPI&cluster=6809.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CSOD1070DE08QPI.

FEATURES
Source Location/Qualifiers
1..1149
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1070YJ16"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 240 a 285 c 285 g 298 t 41 others

ORIGIN

Query Match 34.0%; Score 773; DB 13; Length 1149;
Best Local Similarity 100.0%; Pred. No. 1,8e-123;
Matches 773; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 GAGGGGAGCAGAGGGGAGTGTGTGGCGCTGCGGGGGTTCTGCAGTGGGGTTTCCAGGAG 152

Db 123 GAGGGGAGCAGAGGGGAGTGTGTGGCGCTGCGGGGGTTCTGCAGTGGGGTTTCCAGGAG 182

QY 153 CTGAGCTTAAAGAGTGGGAGCGTCCCTGGGCGCTCAGACAGGCGGCTGGCTGATC 212

Db 183 CTGAGCTTAAAGAGTGGGAGCGTCCCTGGGCGCTCAGACAGGCGGCTGGCTGATC 242

QY 213 ATCTCCATCTTCTGCGGTACCCCTTGTGTTTATCGCATTCCTTTCTACAG 272

Db 243 ATCTCCATCTTCTGCGGTACCCCTTGTGTTTATCGCATTCCTTTCTACAG 302

QY 273 GAGACTACTCTATCCACTCTTTCCATACCTTTACAGGCTCTCAATTCCTATTATTAC 332

Db 303 GAGACTACTCTATCCACTCTTTCCATACCTTTACAGGCTCTCAATTCCTATTATTAC 362

QY 333 TTGGAAACAGAGTCTACAGCTCCCTGTGTGATGTGCTTCAGTTCCTACCTTGA 392

Db 363 TTGGAAACAGAGTCTACAGCTCCCTGTGTGATGTGCTTCAGTTCCTACCTTGA 422

QY 393 CTATATGGGCGGACACATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTACCT 452

Db 423 CTATATGGGCGGACACATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTACCT 482

QY 453 CTGCTGATATCTATTAACCTGACCGGCACTACGATATCAATGTGCAATGCCACAT 512

Db 483 CTGCTGATATCTATTAACCTGACCGGCACTACGATATCAATGTGCAATGCCACAT 542

QY 513 TGTGTTCTGACTTGAAGCTGATGTTGGCTGTGACTCTTGAAGGAGGAAAGAT 572

Db 543 TGTGTTCTGACTTGAAGCTGATGTTGGCTGTGACTCTTGAAGGAGGAAAGAT 602

QY 573 CAGAAATTCCTGTCTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCCCTGCTG 632

Db 603 CAGAAATTCCTGTCTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCCCTGCTG 662

QY 633 GAAGTTGCTGTTTCTCTCTACTTCTATGAGGCTTTGTAGGCGCCCAAGTTCTCAATG 692

Db 663 GAAGTTGCTGTTTCTCTCTACTTCTATGAGGCTTTGTAGGCGCCCAAGTTCTCAATG 722

QY 693 AATCATCATGAGAGGTGTGAGGAGAGAGTATGATCATACAGAGAAATACCAAC 752

Db 723 AATCATCATGAGAGGTGTGAGGAGAGAGTATGATCATACAGAGAAATACCAAC 782

QY 753 AGCATATTCCTGCTCTCAAGCGCTGAGTGTGGCTTTTCTACTAGTGGCTACACA 812

Db 783 AGCATATTCCTGCTCTCAAGCGCTGAGTGTGGCTTTTCTACTAGTGGCTACACA 842

QY 813 CTGCTAGGCGCCCAATCAGAGAGCTATCTCTCATGAGAGCTATGACAA 865

Db 843 CTGCTAGGCGCCCAATCAGAGAGCTATCTCTCATGAGAGCTATGACAA 895

RESULT 5
BQ917856 974 bp mRNA linear EST 20-AUG-2002
LOCUS BQ917856
DEFINITION AGENCOUNT 8801447 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6311255
5', mRNA sequence.
ACCESSION BQ917856
VERSION BQ917856.1 GI:22332554
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 974)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: Reggen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov/plate: LLNL13734 row: n column: 24
High quality sequence scop: 755.
Location/Qualifiers
1..974
/organism="Mus musculus"
/mol_type="mRNA"


```

/db_xref="taxon:10090"
/clone="IMAGE:6311255"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_129"
/Note="Organ: olfactory epithelium; Vector:
pMV-SpOrf6.1.ccd; Site_1: EcoRV; Site_2: NotI; Cloned
unidirectionally. Primer: Oligo dt. Average insert size
2.2 kb. Constructed by Reegen, Invitrogen Corp. Note: this
is a NIH_MGC library."
BASE COUNT      199 a      260 c      248 g      267 t
ORIGIN

```

```

Query Match      33.7%; Score 767; DB 13; Length 974;
Best Local Similarity 100.0%; Pred. No. 2.1e-122;
Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 68 CCGAGTTAAGATGGCGCTCTCAGCGGAGGGAAGGAGGAGCTGTGGCGCTGGCGG 127
DB 36 CCGAGTTAAGATGGCGCTCTCAGCGGAGGGAAGGAGGAGCTGTGGCGCTGGCGG 95
QY 128 GGTTCGAGTGGGGTTCCAGAGCTGAGCTTAACTGAGTGGGAGCGCTGGCGG 187
DB 96 GATTTCGAGTGGGGTTCCAGAGCTGAGCTTAACTGAGTGGGAGCGCTGGCGG 155
QY 188 GTCAAGACAGGCGCTGCGCTGATCATCTTCATCTTCAGGTTACCTTTGCTTTGT 247
DB 156 GTCAAGACAGGCGCTGCGCTGATCATCTTCATCTTCAGGTTACCTTTGCTTTGT 215
QY 248 TTATGGGCTTACCTTTTCAAGAGAACCTTACCTTACCTTTCATACCTTTAC 307
DB 216 TTATGGGCTTACCTTTTCAAGAGAACCTTACCTTACCTTTCATACCTTTAC 275
QY 308 AGGCGCTCAATGCTTATTTAATCTTGGAAACAGCTTACCTTCCCTGCTGTAT 367
DB 276 AGGCGCTCAATGCTTATTTAATCTTGGAAACAGCTTACCTTCCCTGCTGTAT 335
QY 368 TGTGCTTCAAGTCTTCAATCTTGAATAAGGCGGACATCATCTGCGCTCTACTAC 427
DB 336 TGTGCTTCAAGTCTTCAATCTTGAATAAGGCGGACATCATCTGCGCTCTACTAC 395
QY 428 CTTTTCCTTCCAGATGGCTTACCTTGTGCTGGATGATCTTACCTGCGCAACTA 487
DB 396 CTTTTCCTTCCAGATGGCTTACCTTGTGCTGGATGATCTTACCTGCGCAACTA 455
QY 488 CGATATCAAGTGGAGCAATGCAATGATGCTGATGATGATGATGATGATGATGAT 547
DB 456 CGATATCAAGTGGAGCAATGCAATGATGCTGATGATGATGATGATGATGATGAT 515
QY 548 TGAATCACTTGAAGGAGGAAAGATCAGAAATTCCTTCTGAGCAACAGAAATATGC 607
DB 516 TGAATCACTTGAAGGAGGAAAGATCAGAAATTCCTTCTGAGCAACAGAAATATGC 575
QY 608 CATACGTGTGTTCTTCCCTGCTGGAAGTGTGCTTCTCTACTTATGGGCTT 667
DB 576 CATACGTGTGTTCTTCCCTGCTGGAAGTGTGCTTCTCTACTTATGGGCTT 635
QY 668 CTTGAGGAGGCGGAGCTTCAATGATGATGATGATGATGATGATGATGATGAT 727
DB 636 CTTGAGGAGGCGGAGCTTCAATGATGATGATGATGATGATGATGATGATGAT 695
QY 728 TGAATCACTTGAAGGAGGAAAGATCAGAAATTCCTTCTGAGCAACAGAAATATGC 787
DB 696 TGAATCACTTGAAGGAGGAAAGATCAGAAATTCCTTCTGAGCAACAGAAATATGC 755
QY 788 CTTTTCCTTCCAGATGGCTTACCTTGTGCTGGATGATCTTACCTGCGCAACTA 834
DB 756 CTTTTCCTTCCAGATGGCTTACCTTGTGCTGGATGATCTTACCTGCGCAACTA 802

```

```

RESULT 5
BUS41837      BUS41837      916 bp      mRNA      linear      EST 13-SEP-2002
LOCUS
DEFINITION  AGENCOURT_10325398 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6573588

```

```

ACCESSION      5', mRNA sequence.
BUS41837
VERSION      BUS41837.1 GI:22852320
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
NIH-MGC http://mhc.nci.nih.gov/
1 (bases 1 to 916)
AUTHORS
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bim-remail.nih.gov
Tissue Procurement: DCTD/DRP

```

```

FEATURES
source
1..916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6573588"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_40"
/Note="Organ: prostate; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."
BASE COUNT      189 a      249 c      221 g      257 t
ORIGIN

```

```

Query Match      33.0%; Score 751; DB 13; Length 916;
Best Local Similarity 100.0%; Pred. No. 1.2e-119;
Matches 751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 CTCACCGAGGAGGAGGAGGAGGAGCTGTGGCGCTGGCGGAGTTCGACATCGGTTT 145
DB 17 CTCACCGAGGAGGAGGAGGAGGAGGAGCTGTGGCGCTGGCGGAGTTCGACATCGGTTT 76
QY 146 CCGAGAGCTGAGCTTAAACAGTGGCGAGCTCCCTGGGCGCTGAGAAACAGGCGCTGG 205
DB 77 CCGAGAGCTGAGCTTAAACAGTGGCGAGCTCCCTGGGCGCTGAGAAACAGGCGCTGG 136
QY 206 GCTGATCATCTTCAATCTTCCGAGTACCTTGTGATGATGATGATGATGATGATGAT 265
DB 137 GCTGATCATCTTCAATCTTCCGAGTACCTTGTGATGATGATGATGATGATGATGAT 196
QY 266 CTACAAGAGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 325
DB 197 CTACAAGAGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 256
QY 326 TTTTAACTTTGGAACACAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 385
DB 257 TTTTAACTTTGGAACACAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 316
QY 386 CTTGAGCTTAATGGGCGGAGACATCATCTGCGCTGCTGATGATGATGATGATGATGAT 445
DB 317 CTTGAGCTTAATGGGCGGAGACATCATCTGCGCTGCTGATGATGATGATGATGATGAT 376
QY 446 CTACCTTGTGCTGATATTAATCACTGCGACCGGCAACTAGATATCAATGAGCAAT 505

```

Db 377 CTACCTTCGCTGATGATCTATTACACTGCCACCGCACTACGATATCAAGTGACAAAT 436

QY 506 GCCACATTTGTTCTGACTTTGAAGCTGATTTGGCTGTGACTACTTTGACGAGG 565

Db 437 GCCACATTTGTTCTGACTTTGAAGCTGATTTGGCTGTGACTACTTTGACGAGG 496

QY 566 GAAAGATCAAAATTCCTTCCTCTGAGCAACAAATATGCCATATGCTGTCTTC 625

Db 497 GAAAGATCAAAATTCCTTCCTCTGAGCAACAAATATGCCATATGCTGTCTTC 556

QY 626 CCTGCTGGAAGTGTGCTGCTTCTCTCACTTCTATGAGGCTCTTGGTAGGCCCCAGTT 685

Db 557 CCTGCTGGAAGTGTGCTGCTTCTCTCACTTCTATGAGGCTCTTGGTAGGCCCCAGTT 616

QY 686 CTCATGATCACTACATGAGGCTGTGAGGAGAGCTGATTTGACATACAGAAAGAT 745

Db 617 CTCATGATCACTACATGAGGCTGTGAGGAGAGCTGATTTGACATACAGAAAGAT 676

QY 746 ACCAAACAGATCATCTTCTGCTCTCAAGCGCTGAGTCTGGCTTTTCTACTAGTGG 805

Db 677 ACCAAACAGATCATCTTCTGCTCTCAAGCGCTGAGTCTGGCTTTTCTACTAGTGG 736

QY 806 CTACACACTGCTCAGCCCCCAGATCAGAA 836

Db 737 CTACACACTGCTCAGCCCCCAGATCAGAA 767

RESULT 7
BG823157 800 bp mRNA linear EST 22-MAY-2001
LOCUS 602726506F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:486169 5',
DEFINITION mRNA sequence.
VERSION BG823157.1 GI:14170744
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 800)
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM1732 row: k column: 02
High quality sequence stop: 776.
Location/Qualifiers
1..800
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:486169"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/lab_host="NIH_MGC_15"
/clone_id="NIH_MGC_15"
/note="Organ: colon; Vector: POT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 161 a 218 c 199 g 222 t

Query Match 33.0%; Score 750; DB 12; Length 800;
Best Local Similarity 100.0%; Pred. No. 2e-119;
Matches 750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GGGGGGTCCTGTCGCGGGGCTCCGGAGTTAAGATGGCGTCTCAGCGGAGGAGGAGG 106

Db 4 GGGGGGTCCTGTCGCGGGGCTCCGGAGTTAAGATGGCGTCTCAGCGGAGGAGGAGG 63

QY 107 GACTGTGTGGGCTGCGCGGGGTTCTGCACTCGGGTTTCAAGAGCTGAGCTTAACA 166

Db 64 GACTGTGTGGGCTGCGCGGGGTTCTGCACTCGGGTTTCAAGAGCTGAGCTTAACA 123

QY 167 GTTGGGAGCGTCCCTGGGCGCGTCAAGACAGCGCTGCGCTGATCTTCCATCTTCT 226

Db 124 GTTGGGAGCGTCCCTGGGCGCGTCAAGACAGCGCTGCGCTGATCTTCCATCTTCT 183

QY 227 GGGTTACCCCTTGTGTTTATATGCGATTACCTTTCTCAAGAGAGCTACCTCAT 286

Db 184 GGGTTACCCCTTGTGTTTATGCGATTACCTTTCTCAAGAGAGCTACCTCAT 243

QY 287 CCACCTTTCATACCTTTTACAGGCTCTCAATTGCTTATTTTGAACAGCT 346

Db 244 CCACCTTTCATACCTTTTACAGGCTCTCAATTGCTTATTTTGAACAGCT 303

QY 347 CTACCACTCCCTGCTGTGATTTGCTTCACTTCTTCCATCTTCCACTAATGGCGCAC 406

Db 304 CTACCACTCCCTGCTGTGATTTGCTTCACTTCTTCCATCTTCCACTAATGGCGCAC 363

QY 407 CATCACTGCGGCTCTCACTACCTTTTGTCCAGATGGCTCACTCTGCGTGGATACGA 466

Db 364 CATCACTGCGGCTCTCACTACCTTTTGTCCAGATGGCTCACTCTGCGTGGATACGA 423

QY 467 TTACACTGCCACCGGCACTACGATATCAAGTGAATGCCAATGTGTTCTGACTT 526

Db 424 TTACACTGCCACCGGCACTACGATATCAAGTGAATGCCAATGTGTTCTGACTT 483

QY 527 GAAGCTGATTTGTTGGCTGTGACTTATCTTGAACGAGGAGAAAGATCAGATTCCTT 586

Db 484 GAAGCTGATTTGTTGGCTGTGACTTATCTTGAACGAGGAGAAAGATCAGATTCCTT 543

QY 587 CTCTGAGCAACAGAAATATGCCATACGATGATCTTCCCTGCTGAGATTCCTGAT 646

Db 544 CTCTGAGCAACAGAAATATGCCATACGATGATCTTCCCTGCTGAGATTCCTGAT 603

QY 647 CTCTGATTTCTATGGGCTCTTCTGTGAGGCGCCAGTTCTCAATGATCTACTAGAA 706

Db 604 CTCTGATTTCTATGGGCTCTTCTGTGAGGCGCCAGTTCTCAATGATCTACTAGAA 663

QY 707 GCTGTGCAAGGAGAGCTGATTTGACATACCAAGAAATATCCAAACAGATCTTCTGC 766

Db 664 GCTGTGCAAGGAGAGCTGATTTGACATACCAAGAAATATCCAAACAGATCTTCTGC 723

QY 767 TCTCAAGCGCTGAGTCTGGGCTTTTCTA 796

Db 724 TCTCAAGCGCTGAGTCTGGGCTTTTCTA 753

RESULT 8
CA489207 911 bp mRNA linear EST 14-NOV-2002
LOCUS CA489207
DEFINITION AGENCOURT_10809550 MAPCL Homo sapiens cDNA clone IMAGE:6721545 5',
VERSION CA489207.1 GI:24951998
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 911)
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL
COMMENT

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL4282 row: n column: 09
High quality sequence stop: 657.
Location/Qualifiers

FEATURES
source

1. 911
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6721545"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1
/lncap"
/lab_host="EMD10B"
/clone_1lb="MAPL"
/note="Vector: pCMV-Sport6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dt. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkok Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

BASE COUNT 219 a 253 c 203 g 234 t 2 others

Query Match 32.9%; Score 747; DB 14; Length 911;

Best Local Similarity 100.0%; Pred. No. 5.7e-119; Indels 0; Gaps 0;

Matches 747; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

764 TCGCTCAAGCGCTGAGTCTGGGCTTTCTACTAGTGGGTACACACTGCTGAGCC 823

1 TCGCTCAAGCGCTGAGTCTGGGCTTTCTACTAGTGGGTACACACTGCTGAGCC 60

824 CCACATCACAAGAACTATCTCTCACTGAAGATGAACAACACCCCTTCTGTTCCG 883

61 CCACATCACAAGAACTATCTCTCACTGAAGATGAACAACACCCCTTCTGTTCCG 120

884 CTGCATGATCATGCTGATCTGGGGCAAGTTGTGCTGTAACAATATGCACTGTGGCT 943

121 CTGCATGATCATGCTGATCTGGGGCAAGTTGTGCTGTAACAATATGCACTGTGGCT 180

944 GGTCAACAAGAGATGATGATTTTGAAGGCTGAGCTTCAATGCTTTGAAGAAAGG 1003

181 GGTCAACAAGAGATGATGATTTTGAAGGCTGAGCTTCAATGCTTTGAAGAAAGG 240

1004 CAAGCAAGTGGAGTCTGTGTCACCAATGAAGGTGTGCTTTGAAGAAAGG 1063

241 CAAGCAAGTGGAGTCTGTGTCACCAATGAAGGTGTGCTTTGAAGAAAGG 300

1064 CTTCAGTGCACATGCTCATTTCAATCAACACCAAGCTGGGTGGCCGCTACAT 1123

301 CTTCAGTGCACATGCTCATTTCAATCAACACCAAGCTGGGTGGCCGCTACAT 360

1124 CTTCAGTGCACATGCTCATTTCAATCAACACCAAGCTGGGTGGCCGCTACAT 1183

361 CTTCAGTGCACATGCTCATTTCAATCAACACCAAGCTGGGTGGCCGCTACAT 1243

1184 CTTCAGTGCACATGCTCATTTCAATCAACACCAAGCTGGGTGGCCGCTACAT 1243

421 CTTCAGTGCACATGCTCATTTCAATCAACACCAAGCTGGGTGGCCGCTACAT 480

1244 CATTTGATTTGGAAGAAGAGGCTGCCAGGCTCATTTCAAGAGAGGCTGAGCA 1303

481 CATTTGATTTGGAAGAAGAGGCTGCCAGGCTCATTTCAAGAGAGGCTGAGCA 540

Qy 1304 GCTGCGCGCATTAATCTGCTCTCCAGCCCTTCTACTATTGTTGTAACAAGACCATCAGT 1363

Db 541 GCTGCGCGCATTAATCTGCTCTCCAGCCCTTCTACTATTGTTGTAACAAGACCATCAGT 600

Qy 1364 GCTTTTCATGGTTACTCCATGATGCTGCTTCTGCTCTTCACTGAGGAGCAATGCTTAA 1423

Db 601 GCTTTTCATGGTTACTCCATGATGCTGCTTCTGCTCTTCACTGAGGAGCAATGCTTAA 660

Qy 1424 GGTATTAATATCACTATTTCTGTTGGCAGATCTTCTTCTGAGCTTATCATATT 1483

Db 661 GGTATTAATATCACTATTTCTGTTGGCAGATCTTCTTCTGAGCTTATCATATT 720

Qy 1484 GCTTATATTCACAAAGCAATGTGCC 1510

Db 721 GCTTATATTCACAAAGCAATGTGCC 747

RESULT 9

BM557200 1053 bp mRNA linear EST 20-FEB-2002

LOCUS BM557200

DEFINITION AGENCOURT 6579051 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466619

5' mRNA sequence.

ACCESSION BM557200

VERSION BM557200.1 GI:18799007

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/.

1 (bases 1 to 1053)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM1967 row: e column: 20
High quality sequence stop: 677.
Location/Qualifiers

FEATURES
source

1. 1053
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5466619"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

BASE COUNT 230 a 313 c 237 g 267 t 6 others

Query Match 32.6%; Score 741; DB 12; Length 1053;

Best Local Similarity 99.8%; Pred. No. 5.4e-118; Indels 0; Gaps 0;

Matches 841; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1092 ATCAACACCAAGCGCTGGGCTGCGCTACATCTTCAACGCTCAAGTCTTGAAT 1151

Db 1 ATCAACACCAAGCGCTGGGCTGCGCTACATCTTCAACGCTCAAGTCTTGAAT 60

1152 AAGAACTCTCTCAGGGTCTCTGTTGCTATCTGCGCTGCGAGGCTGCACTCA 1211

Qy	648	TCCTACTCTATGGGGGCGCTTCTGGTAGGGCCCCGATTCTCAATTAATCAGTACATCTAG	707
Db	632	TCCTACTCTATGGGGGCGCTTCTGGTAGGGCCCCGATTCTCAATTAATCAGTACATCTAG	691
Qy	708	CTGTGTGACGAGGAGCTGATTTGACATACACAGGAAGATACCAAAACAGCATCTCTGCT	767
Db	692	CTGTGTGACGAGGAGCTGATTTGACATACACAGGAAGATACCAAAACAGCATCTCTGCT	751
Qy	768	CTCAAGCGCCCTGA	780
Db	752	CTCAAGCGCCCTGA	764
RESULT 11			
LOCUS	BG753831	887 bp	mRNA
DEFINITION	602733053P1 NIH_MGC_43 Homo sapiens CDNA clone IMAGE:4876525 5',		
ACCESSION	BG753831		
VERSION	BG753831.1	GI:14064484	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 887)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
Unpublished	Contact: Robert Strusberg, Ph.D.		
Email: cgsphs-remail.nih.gov			
Tissue Procurement: ATCC			
CDNA Library Preparation: Ling Hong/Rubin Laboratory			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNI)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LILNI at:			
http://image.llnl.gov			
Plate: LILNI1759	row: j	column: 14	
High quality sequence stop: 814.			
Location/Qualifiers			
1. 887			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4876525"			
/tissue_type="normal pigmented retinal epithelium"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_43"			
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:			
EcoRI; CDNA made by oligo-dT priming. Directionally			
cloned into EcoRI/XhoI sites using the following 5'			
adaptor: GGCACGAG(G). Library constructed by Ling Hong			
in the laboratory of Gerald M. Rubin (University of			
California, Berkeley) using ZAP-cDNA synthesis kit			
(Stratagene) and Superscript II RT (Life Technologies).			
Note: this is a NIH_MGC library. "			
BASE COUNT	208 a	203 g	226 t
ORIGIN			
Query Match	31.5%	Score 715;	DB 10; Length 887;
Best Local Similarity	99.9%;	Pred. No. 1.7e-11;	
Matches 765; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1020	GCGTGTGCAACATGAAGGTGTGGCTCTTTGAACAACACCCCGCTTACTGAGCACTT	1079
Db	1	GCGTGTGCAACATGAAGGTGTGGCTCTTTGAACAACACCCCGCTTACTGAGCACTT	60
Qy	1080	GCGTCAATTCACATCAACACCAACGCGCTGGTGGCCGCGCTACATCTTCAACAGACTCAAG	1139
Db	61	GCGTCAATTCACATCAACACCAACGCGCTGGTGGCCGCGCTACATCTTCAACAGACTCAAG	120
Qy	1140	TTCCCTGGAATTAAGAACTCTCTCAGAGGTCTCTGCTTGTAATTCGAGCCCTCTGCGAC	1199

Db	121	TTCTTGGAAATTAAGAACTCTCAGAGGTTCTGTTGCTATTTCCTGAGCCCTTGAC	180
QY	1200	GGCCGTGACCTCAGGATACCTGGTCTGCTTCAGATGGAATTCCTCATTTGTAATGTGAA	1259
Db	181	GGCCGTGACCTCAGGATACCTGGTCTGCTTCAGATGGAATTCCTCATTTGTAATGTGAA	240
QY	1280	AGACAGGCTGCAGGCTCATTCAGAGAGCCCACTGAGCAGCTGGCCGCAATTA	1319
Db	241	AGACAGGCTGCAGGCTCATTCAGAGAGCCCACTGAGCAGCTGGCCGCAATTA	300
QY	1320	GTCCTCCAGCCCTCTCTATTTGGTGCACAGACCATTCACCTGGCTTCATGGGTAC	1379
Db	301	GTCCTCCAGCCCTCTCTATTTGGTGCACAGACCATTCACCTGGCTTCATGGGTAC	360
QY	1380	TCCATGACCTGCTCTGCTCTTCACCTGGGACAAATGAGCTTAAGGTATTAATCCATC	1439
Db	361	TCCATGACCTGCTCTGCTCTTCACCTGGGACAAATGAGCTTAAGGTATTAATCCATC	420
QY	1440	TATTTCTCTGGCCATCTTTCTCTGAGCCCTATTCATATTCCTTATATTCACAAA	1499
Db	421	TATTTCTCTGGCCATCTTTCTCTGAGCCCTATTCATATTCCTTATATTCACAAA	480
QY	1500	GCAATGGTCCAAAGAAAGAACTTAAAGAAATGAAATTCATATTCCTTATATTCACAAA	1559
Db	481	GCAATGGTCCAAAGAAAGAACTTAAAGAAATGAAATTCATATTCCTTATATTCACAAA	540
QY	1560	TGTGCGGAGCTGTGTCAGAAATCTAGTCTCCCTTTTCACAGCACTCTTTGCCCCAGA	1619
Db	541	TGTGCGGAGCTGTGTCAGAAATCTAGTCTCCCTTTTCACAGCACTCTTTGCCCCAGA	600
QY	1620	GCAGAGATGAGAAAGCCAGGAGGTGAGAGATGCATTCAGCTGTGCTCTGTGTC	1679
Db	601	GCAGAGATGAGAAAGCCAGGAGGTGAGAGATGCATTCAGCTGTGCTCTGTGTC	660
QY	1680	CAGCAGATCTTCACTTTGGGGCCAAAGGGAACTTTTGTGAGAAAGGCTCTTGTCTT	1739
Db	661	CAGCAGATCTTCACTTTGGGGCCAAAGGGAACTTTTGTGAGAAAGGCTCTTGTCTT	720
QY	1740	TGTACACCACGCTGGAATGCAGTGGCGGAGATCTCAGCTCACCCGCA	1785
Db	721	TGTACACCACGCTGGAATGCAGTGGCGGAGATCTCAGCTCACCCGCA	766
RESULT 12			
Bi760522			
LOCUS			
DEFINITION	603045066F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185120 5',		
ACCESSION	Bi760522		
VERSION	Bi760522.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 847)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-ri@mail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LINL at:		
	http://image.llnl.gov		
	Plate: L14M11462 row: d column: 17		
	High quality sequence stop: 807.		
FEATURES			
source	1..847		
	Location/Qualifiers		

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5185120"
/lab_host="DH10B"
/clone_1ib="NIH MGC 116"
/note="Organ: pooled colon, kidney, stomach; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."
BASE COUNT      170 a      236 c      210 g      231 t
ORIGIN

```

```

Query Match      31.3%; Score 711; DB 12; Length 847;
Best Local Similarity 100.0%; Pred. No. 8.4e-113;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

47 GGGGGGCTCTGGGGGCTCCGGAGTTAGTGGCTCTCCAGCGGAGGGGACGAGG 106
17 GGGGGGCTCTGGGGGCTCCGGAGTTAGTGGCTCTCCAGCGGAGGGGACGAGG 76
107 GACTGTGTGGGCGCTGGGGGCTTCGAGTCGGGTTTCAGAGGCTGAGCTTAACA 166
77 GACTGTGTGGGCGCTGGGGGCTTCGAGTCGGGTTTCAGAGGCTGAGCTTAACA 136
167 GTTGGCGACGTCCCTGGGCGCTGACAGACGGCGCTGGCTGATCTTCATCTTCT 226
137 GTTGGCGACGTCCCTGGGCGCTGACAGACGGCGCTGGCTGATCTTCATCTTCT 196
227 GGGTTACCCCTTGTGTTTATGGGCAATACCTTTTCAAGAGAGCACTACCTCAT 286
197 GGGTTACCCCTTGTGTTTATGGGCAATACCTTTTCAAGAGAGCACTACCTCAT 256
287 CCACCTTTCATACCTTTTACAGGCGCTCAATGCTTATTTTAACTTTGAAACAC 346
257 CCACCTTTCATACCTTTTACAGGCGCTCAATGCTTATTTTAACTTTGAAACAC 316
347 CTACCACTCCCTGTGTGATGTGCTTCAAGTTCCTCATCTTCGACTTAATGGGCG 406
317 CTACCACTCCCTGTGTGATGTGCTTCAAGTTCCTCATCTTCGACTTAATGGGCG 376
407 CATCAGCGGCTCTCACTACCTTTTGCCTTCCAGATGGCCCTACCTTCTGGCTGAT 466
377 CATCAGCGGCTCTCACTACCTTTTGCCTTCCAGATGGCCCTACCTTCTGGCTGAT 436
467 TTACACTGCCAGCGCAACTACGATATCAAGTGGACATGCCAATGTGTTCGACTTT 526
437 TTACACTGCCAGCGCAACTACGATATCAAGTGGACATGCCAATGTGTTCGACTTT 496
527 GAAAGTGTGTTGGCTGTGACTTACCTTTGACGAGGAGGAAAGATCAGAATTCCTT 586
497 GAAAGTGTGTTGGCTGTGACTTACCTTTGACGAGGAGGAAAGATCAGAATTCCTT 556
587 CTCTGAGCAACAGAAATATGCAATAGTGGTTCCTCCCTGTGGAAGTGTGGTTT 646
557 CTCTGAGCAACAGAAATATGCAATAGTGGTTCCTCCCTGTGGAAGTGTGGTTT 616
647 CTCTACTTCTATGGGCTTTCTTGTAGGGCCCACTTCTCAATGATCACTACATGAA 706
617 CTCTACTTCTATGGGCTTTCTTGTAGGGCCCACTTCTCAATGATCACTACATGAA 676
707 GCTGTGTGAGGAGGAGTGTGATGACATACGAGAAATACCAACAGCAT 757
677 GCTGTGTGAGGAGGAGTGTGATGACATACGAGAAATACCAACAGCAT 727
Db
Qy

```

```

RESULT 13
AL553069      1201 bp      mRNA      linear      EST 31-MAY-2003
LOCUS
DEFINITION
AL553069 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1072Y019 5-PRIME, mRNA sequence.
ACCESSION
AL553069
VERSION
AL553069.2 GI:31274883
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12892559.
COMMENT
Genoscope - Centre National de Sequenage
BP 191 91006 EVRY cedex - France
Email: segr@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6809.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1072AH10P1&cluster=6809.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1072AH10P1.
Location/Qualifiers
1..1201

```

FEATURES

```

source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1072Y019"
/issue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      288 a      273 c      274 g      309 t      57 others
ORIGIN

```

```

Query Match      31.1%; Score 708; DB 9; Length 1201;
Best Local Similarity 100.0%; Pred. No. 2.1e-112;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

512 TTGTGTTGACCTTGAAGCTGATGTTGGTGTGACTTATGACGAGGAAAGA 571
63 TTGTGTTGACCTTGAAGCTGATGTTGGTGTGACTTATGACGAGGAAAGA 122
572 TCAGAAATTCCTTGTCTCTGAGCAAGAAATATGCAATCGTGTTCCTTCCCTGCT 631
123 TCAGAAATTCCTTGTCTCTGAGCAAGAAATATGCAATCGTGTTCCTTCCCTGCT 182
632 GGAAGTGTGTTTCTCTACTTCTATGAGGCGCTTCTTGTGAGGCGCCCACTTCAAT 691
183 GGAAGTGTGTTTCTCTACTTCTATGAGGCGCTTCTTGTGAGGCGCCCACTTCAAT 242
692 GAATCACTATGATGAGGTGTGAGGAGAGCTGATGACATACAGAAAGATACCAAA 751
243 GAATCACTATGATGAGGTGTGAGGAGAGCTGATGACATACAGAAAGATACCAAA 302
752 CAGCATCATTCCTGCTCTCAAGCGCTGAGTCTGGGCTTTTCTACTAGTGGCTACAC 811
303 CAGCATCATTCCTGCTCTCAAGCGCTGAGTCTGGGCTTTTCTACTAGTGGCTACAC 362
812 ACTGCTCAGGCCCACTACAGAAAGCTATCTCTCACTGAAAGACTATATCAACACCC 871
363 ACTGCTCAGGCCCACTACAGAAAGCTATCTCTCACTGAAAGACTATATCAACACCC 422
872 CTTGTGTTCCGCTGATGATGATGCTGATCTGGGCAAGTTGTGCTGTATCAATATGT 931
Qy
Db

```

|||||
423 CTTGCTGTCCTCCCTGATGATGATGCTGAGGCAAGTTTGCTTACAAATATGT 482
QY 932 CACCTGTTGGCTGTGACAGAGAGATGATGATTTTGAAGGCTGGCTTCAATGGCTT 991
Db 483 CACCTGTTGGCTGTGACAGAGAGATGATGATTTTGAAGGCTGGCTTCAATGGCTT 542
QY 992 TGAAGAAAAGGCAAGGCAAGAGTGGATGCTGTGCAACATGAAGTGTGCTTTTGA 1051
Db 543 TGAAGAAAAGGCAAGGCAAGAGTGGATGCTGTGCAACATGAAGTGTGCTTTTGA 602
QY 1052 AACAAACCCCGCTTCACTGCGACCATGCTCTTCAATCAATCAACACGCTGGGT 1111
Db 603 AACAAACCCCGCTTCACTGCGACCATGCTCTTCAATCAATCAACACGCTGGGT 662
QY 1112 GGCCGCTGATCTTCAAGCACTCAAGTCTTGTGAATAAAGATCTCTCAGGCTCT 1171
Db 663 GGCCGCTGATCTTCAAGCACTCAAGTCTTGTGAATAAAGATCTCTCAGGCTCT 722
QY 1172 CTCGTTGCTATTCCTGCGCTCTGCGACGCGCTGCACTCAGATACCT 1219
Db 723 CTCGTTGCTATTCCTGCGCTCTGCGACGCGCTGCACTCAGATACCT 770

RESULT 14
LOCUS BM553135 1097 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6572558 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467062
5', mRNA sequence.
ACCESSION BM553135
VERSION BM553135.1 GI:18791600
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1097)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1968 row: h column: 07
High quality sequence stop: 666.
Location/Qualifiers
1..1097
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5467062"
/feature_type="amelanotic melanoma, cell line"
/lab_host="NIH MGC 41"
/note="Organ: skin; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

BASE COUNT 244 a 315 c 254 g 275 t 9 others
ORIGIN

Query Match 31.1%; Score 707; DB 12; Length 1097;
Best Local Similarity 100.0%; Pred. No. 3.3e-112;

Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1092 ATCAACACCAACGCGCTGGGTGGCCCGCTACATCTTCAAGCACTCAAGTCTTGGAAAT 1151
Db 1 ATCAACACCAACGCGCTGGGTGGCCCGCTACATCTTCAAGCACTCAAGTCTTGGAAAT 60
QY 1152 AAGAAGCTCTCAGAGGTCTCTGTTGATATTCCTGAGCCCTTGGCAGGCTGACATCA 1211
Db 61 AAGAAGCTCTCAGAGGTCTCTGTTGATATTCCTGAGCCCTTGGCAGGCTGACATCA 120
QY 1212 GATACCTGCTGCTTCCAGATGGAATTCCTCATTTGTTATTTGGAAGACAGCTGCTC 1271
Db 121 GATACCTGCTGCTTCCAGATGGAATTCCTCATTTGTTATTTGGAAGACAGCTGCTC 180
QY 1272 AGGCTCATTCAGAGAGCCCAACCTTGAGCAAGCTGGCCGCAATTAAGTCTCTCAGGCC 1331
Db 181 AGGCTCATTCAGAGAGCCCAACCTTGAGCAAGCTGGCCGCAATTAAGTCTCTCAGGCC 240
QY 1332 TTCTACTATTTGGTGCACAGACATCCACTGCTCTTCATGAGTTACTCATGACTGCC 1391
Db 241 TTCTACTATTTGGTGCACAGACATCCACTGCTCTTCATGAGTTACTCATGACTGCC 300
QY 1392 TTCTGCTCTTTCACGCGGACCAATGCTTAAAGTGTATTAATTCATCTATTTCTTGCC 1451
Db 301 TTCTGCTCTTTCACGCGGACCAATGCTTAAAGTGTATTAATTCATCTATTTCTTGCC 360
QY 1452 CACATCTTCTTCCCTAGGCTTACTATTCATTTGCTTATATTCACAAAGCATGCTGCCA 1511
Db 361 CACATCTTCTTCCCTAGGCTTACTATTCATTTGCTTATATTCACAAAGCATGCTGCCA 420
QY 1512 AGGAAGAGAAATTAAGAGATGGAATTAATTCATTTCCCTGAGGCTGCTGGGACTG 1571
Db 421 AGGAAGAGAAATTAAGAGATGGAATTAATTCATTTCCCTGAGGCTGCTGGGACTG 480
QY 1572 GTGCAAAATCTACTGCTCTCTCTTTCACAGCACTCTTTGCCCGACAGACAGAAATGA 1631
Db 481 GTGCAAAATCTACTGCTCTCTCTTTCACAGCACTCTTTGCCCGACAGACAGAAATGA 540
QY 1632 AAGCCAGGAGAGTGAAGATGATGATGCTTCCAGCTGTGCTGCTGCGACCAACTCTT 1691
Db 541 AAGCCAGGAGAGTGAAGATGATGATGCTTCCAGCTGTGCTGCTGCGACCAACTCTT 600
QY 1692 CATTGGGGGCCAAAGGGGAACTTTTGTGAGAAGGCGTCTGCTTGTGACCAACGC 1751
Db 601 CATTGGGGGCCAAAGGGGAACTTTTGTGAGAAGGCGTCTGCTTGTGACCAACGC 660
QY 1752 TGAATGCAAGTGGCGGATCTCAGCTCACCGCACTCCACCTCTCG 1798
Db 661 TGAATGCAAGTGGCGGATCTCAGCTCACCGCACTCCACCTCTCG 707

RESULT 15
LOCUS BU597281 827 bp mRNA linear EST 20-SEP-2002
DEFINITION AGENCOURT_8966025 NIH_MGC_142 Homo sapiens cDNA clone IMAGE:6454096
5', mRNA sequence.
ACCESSION BU597281
VERSION BU597281.1 GI:23249040
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 827)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCM2630 row: j column: 17
High quality sequence stop: 559.
Location/Qualifiers

FEATURES

source

```

1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6454096"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_142"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcagcc); Site 2: SfiI (ggcgctcgccg); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAAGCAGAGTGGCCATTAGCGCGG-3' and
5'-ATTCTAGAGCCGAGCGCGCGACATG-dr(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >0.5 kb size fraction (other fractions present in NIH_MGC_141). Library created in the laboratory of M. Brownstein (NIH, NIH). Note: this is a NIH MGC Library."
BASE COUNT      185 a      238 c      200 g      201 t      3 others
ORIGIN

```

Query Match 30.9%; Score 703; DB 13; Length 827;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 703; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1248 GTTATTGTGAAAGACAGGCTGCCAGGCTATTCAAGAGAGCCCGCTGAGCAAGCTG 1307
DB 3 GTTATTGTGAAAGACAGGCTGCCAGGCTATTCAAGAGAGCCCGCTGAGCAAGCTG 62
QY 1308 GCCGCATTACTGCTCCAGCCCTTACTATTGTGTGACAGACCATCACTGGCTC 1367
DB 63 GCCGCATTACTGCTCCAGCCCTTACTATTGTGTGACAGACCATCACTGGCTC 122
QY 1368 TTGATGGGTACTGCATGACTGCTTCTGCTCTTCAAGTGGACAATGGCTTAAGTG 1427
DB 123 TTGATGGGTACTGCATGACTGCTTCTGCTCTTCAAGTGGACAATGGCTTAAGTG 182
QY 1428 TATTAATCCATCTATTCTTGGCCACATCTTCTCTGAGCTTACTATTGCTT 1487
DB 183 TATTAATCCATCTATTCTTGGCCACATCTTCTCTGAGCTTACTATTGCTT 242
QY 1488 TATATTCAAAAGCAATGTGTGCCAAAGAAAGAGATTAAAGAGATGAATATCCATT 1547
DB 243 TATATTCAAAAGCAATGTGTGCCAAAGAAAGAGATTAAAGAGATGAATATCCATT 302
QY 1548 TCCCTGTGTGCTGTGTGGGGAATGTGTGCAGAAACTACTGCTCTCTTTTCAAGCACTC 1607
DB 303 TCCCTGTGTGCTGTGTGGGGAATGTGTGCAGAAACTACTGCTCTCTTTTCAAGCACTC 362
QY 1608 CTTTGGCCCGAGAGAGAAATGAAAAAGCCAGAGAGTGAAGATGCATGCTTCCAGCTG 1667
DB 363 CTTTGGCCCGAGAGAGAAATGAAAAAGCCAGAGAGTGAAGATGCATGCTTCCAGCTG 422
QY 1668 TGCGCTGTGCGCAGCAAGCTTCAATTGGGGGCCAAAGGGGAAACTTTTTTTTGAGAA 1727
DB 423 TGCGCTGTGCGCAGCAAGCTTCAATTGGGGGCCAAAGGGGAAACTTTTTTTTGAGAA 482
QY 1728 GGGGCTTTGCTTTGTCAACCAAGCTGGAATGCAATGCGGGATCTCACTCAACCGCAACC 1787
DB 483 GGGGCTTTGCTTTGTCAACCAAGCTGGAATGCAATGCGGGATCTCACTCAACCGCAACC 542

```

```

QY 1788 TCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCCGCCCAAGTAGTGGGAATACAG 1847
DB 543 TCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCCGCCCAAGTAGTGGGAATACAG 602
QY 1848 GCACGCCACCATGCGCAGCTAATTTTGTATTTCAGTAGAAACGGGATTTGACCAAGTT 1907
DB 603 GCACGCCACCATGCGCAGCTAATTTTGTATTTCAGTAGAAACGGGATTTGACCAAGTT 662
QY 1908 GGCAGGCTGTGTCTGAACTCTGACCGCAAGTATCCACCGG 1950
DB 663 GGCAGGCTGTGTCTGAACTCTGACCGCAAGTATCCACCGG 705

```

Search completed: November 22, 2003, 12:02:52
Job time : 4745 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2003, 07:52:31 ; Search time 600 Seconds
(without alignments)
10226.371 Million cell updates/sec

Title: US-09-938-803-26

Perfect score: 2273

Sequence: 1 99999gtgaagcgatcgtt.....tttccaaaaaataaaaaa 2273

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N.Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2273	100.0	2273	22	Human FLEXHT-6 nuc
2	2211	97.3	2264	21	Human FLEXHT-6 nuc
3	2186	96.2	2234	21	Human FLEXHT-6 nuc
4	1842	81.0	1842	21	Human FLEXHT-6 nuc
5	1842	81.0	1842	24	Human FLEXHT-6 nuc
6	1842	81.0	1842	24	Human FLEXHT-6 nuc
7	1461	64.3	1461	22	Human FLEXHT-6 nuc
8	1156	50.9	1867	23	Human FLEXHT-6 nuc

C 9	708	31.1	7461	22	AAS30639	DNA encoding novel
C 10	708	31.1	7461	22	AAS28701	Genomic sequence #
C 11	708	31.1	7461	22	ACA03402	DNA encoding human
C 12	708	31.1	32174	22	ABA15665	Human nervous syst
C 13	708	31.1	32174	22	ABA19477	Human nervous syst
C 14	708	31.1	32174	22	ABA20359	Human nervous syst
C 15	708	31.1	32174	22	ABA21505	Human nervous syst
C 16	708	31.1	32174	22	ABA36280	Human musculoskele
C 17	708	31.1	32174	22	AAS32655	Human genomic DNA
C 18	708	31.1	32174	22	AAS34394	Human DNA for a no
C 19	708	31.1	32174	22	AAS30638	DNA encoding novel
C 20	708	31.1	32174	22	AAL03792	Human reproductive
C 21	708	31.1	32174	22	AAL07447	Human reproductive
C 22	708	31.1	32174	22	AAL07491	Human reproductive
C 23	708	31.1	32174	22	AAS28700	Genomic sequence #
C 24	708	31.1	32174	22	AAI62606	Human breast or ov
C 25	708	31.1	32174	22	AAI62904	Human genomic DNA
C 26	708	31.1	32174	23	ABK72131	Human ovarian anti
C 27	708	31.1	32174	24	ABK91723	Novel ovarian rela
C 28	708	31.1	32174	25	ACA03401	DNA encoding human
C 29	708	31.1	32174	25	ABX59268	cDNA encoding nove
C 30	708	31.1	38771	22	AAK66361	Human immune/haema
C 31	708	31.1	38771	22	AAK66883	Human immune/haema
C 32	708	31.1	38771	22	AAK65604	Human immune/haema
C 33	708	31.1	38771	22	AAK71499	Human immune/haema
C 34	708	31.1	38771	22	AAK72925	Human immune/haema
C 35	708	31.1	38771	22	AAK75870	Human immune/haema
C 36	708	31.1	38771	22	AAK80325	Human immune/haema
C 37	708	31.1	38771	22	AAK83538	Human immune/haema
C 38	708	31.1	38771	25	ABZ74036	Secreted protein g
C 39	708	31.1	38771	25	ABZ74370	Secreted protein g
C 40	708	31.1	38771	25	ABZ74376	Secreted protein g
C 41	708	31.1	38771	25	ABZ67623	Human secreted pro
C 42	708	31.1	38771	25	ABZ67941	Human secreted pro
C 43	708	31.1	38771	25	ABZ67947	Human secreted pro
C 44	708	31.1	222630	24	ABK84349	Human CDNA differe
C 45	596	26.2	1296	23	ABV30301	Human prostate exp

ALIGNMENTS

RESULT 1	
AAAC8075	AAAC8075 standard; cDNA; 2273 BP.
AC	AAAC8075;
XX	
DT	09-MAR-2001 (first entry)
XX	
DE	Human FLEXHT-6 nucleotide sequence SEQ ID NO:61.
XX	
KW	Human; FLEXHT; full-length molecules expressed in human tissue;
KW	diagnosis; gene expression; genetic linkage; genetic variability;
KW	antianemic; anticonvulsant; antiarteriosclerotic; immunomodulatory;
KW	cytostatic; antitumor; antineoplastic; antiproliferative; antidiabetic;
KW	anti-gout; antithyroid; neuroprotective; antiarthritic; osteopathic;
KW	anti-psoriatic; antirheumatic; antileukemic; anemia; gout;
KW	epilepsy; arteriosclerosis; atherosclerosis; developmental disorder;
KW	cancer; immunological disorder; asthma; bronchitis; cirrhosis;
KW	Cronh's disease; diabetes mellitus; Grave's disease; multiple sclerosis;
KW	osteoarthritis; pancreatitis; rheumatoid arthritis; psoriasis;
KW	ulcerative colitis; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200070047-A2.
XX	
PD	23-NOV-2000.
XX	
PF	12-MAY-2000; 2000WO-US13299.
XX	
PR	14-MAY-1999; 99US-0311894.

BR	14-MAY-1999;	98US-0311937.
RR	14-MAY-1999;	98US-0311940.
XX	(INCY-) INCYTE GENOMICS INC.	
PA	Yue H, Tang YT, Lal P, Reddy R, Batra S, Baughn MR, Yang J;	
PI	Azimzai Y, Lu DM, Au-Young J, Shih LH;	
PI	Human FLEKHT protein and DNA sequences, useful for treating	
DR	immunological disorders, developmental disorders, and cancers -	
XX	P-PSDB; AAB36584.	
XX	WPI; 2001-016234/02.	
PS	Claim 5; Page 140; 16pp; English.	
CC	AACG8070 to AAC80124 encode the 55 FLEKHT (full-length molecules	
CC	expressed in human tissues) proteins given in AAB36579 to AAB36633. The	
CC	present invention describes an isolated polypeptide (A) comprising an	
CC	amino acid sequence selected from one of 55 amino acid sequences 42-876	
CC	residues in length, corresponding to FLEKHT-1 to FLEKHT-55, a 90 %	
CC	identical sequence, and a biologically active or immunogenic fragment of	
CC	the sequence. The FLEKHT proteins can have antianemic, anticomaletant,	
CC	antiarteriosclerotic, immunomodulatory, cytostatic, antistematic,	
CC	antiinflammatory, hepatotropic, antidiabetic, anti-gout, antithyroid,	
CC	neuroprotective, antiarthritic, osteopathic, antipsoriatic, antiflcer	
CC	and antineumatic activities, and can be used in gene therapy. The	
CC	polynucleic acid sequences can be used to express the protein sequences.	
CC	Pharmaceutical compositions comprising FLEKHT can be used to treat	
CC	diseases or conditions associated with altered expression of functional	
CC	FLEKHT. The proteins and polynucleotides may be used to diagnose and	
CC	treat disorders including anaemia, epilepsy, arteriosclerosis,	
CC	atherosclerosis, developmental disorders, cancers, and immunological	
CC	disorders such as asthma, bronchitis, cirrhosis, Crohn's disease,	
CC	diabetes mellitus, gout, Grave's disease, multiple sclerosis,	
CC	osteoarthritis, pancreatitis, psoriasis, rheumatoid arthritis, and	
CC	ulcerative colitis.	
XX	Sequence 2273 BP; 494 A; 630 C; 547 G; 602 T; 0 other;	
SQ		
Query Match	100.0%; Score 2273; DB 22; Length 2273;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 2273; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 GGGGGTGAAGCATACTGTTTCGCCGCATTCGCGGCGCGGCACTGGGGGGGTCCCTTG 60	
Dd	1 GGGGGTGAAGCATACGTTTTCCCGCATTCGGGGCGCGCGGACTGGGGGGGTCCCTTG 60	
OY	61 GGGCTCCCGAGTAAGATGCGCTCTTCAGCGAGGGGAGCGAGGGGACTGTGTGGCGC 120	
Dd	61 GGGCTCCCGAGTAAGATGCGCTCTTCAGCGAGGGGAGCGAGGGGACTGTGTGGCGC 120	
OY	121 TGCGCGGGGTTCTGCATGTCGGGTTCCAGAGACTAGCTTAAACAAGTTGGGAGTCCC 180	
Dd	121 TGCGCGGGGTTCTGCATGTCGGGTTCCAGAGACTAGCTTAAACAAGTTGGGAGTCCC 180	
OY	181 TGCGCGCGGTCAAGAACAGGCGCTGCGGCTGATCATCTTCCTGGGTTACCCTTTG 240	
Dd	181 TGCGCGCGGTCAAGAACAGGCGCTGCGGCTGATCATCTTCCTGGGTTACCCTTTG 240	
OY	241 CTTTGTATTATCGGCACTTACCTTTTCTACAGAGAAGCTTACCTTCACTCTTCA 300	
Dd	241 CTTTGTATTATCGGCACTTACCTTTTCTACAGAGAAGCTTACCTTCACTCTTCA 300	
OY	301 CCTTTACAGGCTCTCAATTGCTTATTTAACTTTGAAAACAGCTTACCACTCCCTGC 360	
Dd	301 CCTTTACAGGCTCTCAATTGCTTATTTAACTTTGAAAACAGCTTACCACTCCCTGC 360	
OY	361 TGTGTAATGCTTCACTGCTTCACTCTTCACTGAGCTAATGGGCGCACATCACTGCGTCC 420	
Dd	361 TGTGTAATGCTTCACTGCTTCACTCTTCACTGAGCTAATGGGCGCACATCACTGCGTCC 420	
OY	421 TCACATCTTTGCTTCCAGATGGCCCTTACTTGGCTGGATACATTAACTGTCACGG 480	
Dd	421 TCACATCTTTGCTTCCAGATGGCCCTTACTTGGCTGGATACATTAACTGTCACGG 480	

Db	421	TCACCTACCTTTGCTTCCAGATGCGCTTACCTTCTGCGTGGATCTATTACACTGCGCACCG	480
Qy	481	GCAACTACGATATCAAGTGGACAAATGCGACATGTGTCTGACTTTGAAGCTGATGTT	540
Db	481	GCAACTACGATATCAAGTGGACAAATGCGACATGTGTCTGACTTTGAAGCTGATGTT	540
Qy	541	TGCGCTGTGACTACTTTGACGGAGGGAAAGATCTGAATTCCTTGTCTCTTGACGACACGA	600
Db	541	TGCGCTGTGACTACTTTGACGGAGGGAAAGATCTGAATTCCTTGTCTCTTGACGACACGA	600
Qy	601	AATATGCGCATACGAGGTGTCTTCCCTGCGTGGAAAGTGTGCGTTCCTCTACTTCTATG	660
Db	601	AATATGCGCATACGAGGTGTCTTCCCTGCGTGGAAAGTGTGCGTTCCTCTACTTCTATG	660
Qy	661	GGGCGCTTCTTGATGAGGGCCCCAGTTTCTCAATGAATCTACATGTAAGCTGTGTGACGGAG	720
Db	661	GGGCGCTTCTTGATGAGGGCCCCAGTTTCTCAATGAATCTACATGTAAGCTGTGTGACGGAG	720
Qy	721	AGCTGATTTGACATACCGAGAAAGATACCAACAGCATCATCTCTGCTCTCAAGCGCTTGA	780
Db	721	AGCTGATTTGACATACCGAGAAAGATACCAACAGCATCATCTCTGCTCTCAAGCGCTTGA	780
Qy	781	GTCGAGGCGCTTTTCTACCTAGTGGGCTACACATGCTCAGCGCCGACATCAGAGAAGCT	840
Db	781	GTCGAGGCGCTTTTCTACCTAGTGGGCTACACATGCTCAGCGCCGACATCAGAGAAGCT	840
Qy	841	ATCTCTCACTGAAAGACTATGACAAACACCCCTTCTGGTTCGCGTGCATGTACATGTCTGA	900
Db	841	ATCTCTCACTGAAAGACTATGACAAACACCCCTTCTGGTTCGCGTGCATGTACATGTCTGA	900
Qy	901	TCTGGGGCAGTGTGTGCTGTACAAATATGTCACTGTGTGCTGTCTCAGAGAAGATAT	960
Db	901	TCTGGGGCAGTGTGTGCTGTACAAATATGTCACTGTGTGCTGTCTCAGAGAAGATAT	960
Qy	961	GCAATTTTGA GGGCGCTGGGCTTCAATGGCTTTGAAAGAAAGGGCAAGGCAAGTGGAGATG	1020
Db	961	GCAATTTTGA GGGCGCTGGGCTTCAATGGCTTTGAAAGAAAGGGCAAGGCAAGTGGAGATG	1020
Qy	1021	CCTGTGCGCAATGAGGTGTGTGCTCTTTGAAACAAACCCCGCTTCACTGGACCATTTG	1080
Db	1021	CCTGTGCGCAATGAGGTGTGTGCTCTTTGAAACAAACCCCGCTTCACTGGACCATTTG	1080
Qy	1081	CCTCATTTCAACATCAACCAACGCGCTGGTGGCGCCGCTACATCTTCAACGACTCAAGT	1140
Db	1081	CCTCATTTCAACATCAACCAACGCGCTGGTGGCGCCGCTACATCTTCAACGACTCAAGT	1140
Qy	1141	TCCTTGGAAATTAAGAACTCTCTCAGGGTCTCTGTTGCTATTCCTGGCCCTTGGGACG	1200
Db	1141	TCCTTGGAAATTAAGAACTCTCTCAGGGTCTCTGTTGCTATTCCTGGCCCTTGGGACG	1200
Qy	1201	GCGTGCATCTCAGGATACCTGTGTGCTTCCAGATGGAATTCCTCATTTTATTGGGAAA	1260
Db	1201	GCGTGCATCTCAGGATACCTGTGTGCTTCCAGATGGAATTCCTCATTTTATTGGGAAA	1260
Qy	1261	GACAGGCTGCCAGGCTCAATTCAAGAGAGCCCAACCTTGAAGCAAGCTGGCGCGCATTTACTG	1320
Db	1261	GACAGGCTGCCAGGCTCAATTCAAGAGAGCCCAACCTTGAAGCAAGCTGGCGCGCATTTACTG	1320
Qy	1321	TCCTTCAAGCCTTTCTACTATTTTGTGTGCAACAGACATCTCACTGGGCTCTTCAATGGGTTACT	1380
Db	1321	TCCTTCAAGCCTTTCTACTATTTTGTGTGCAACAGACATCTCACTGGGCTCTTCAATGGGTTACT	1380
Qy	1381	GCATGACAGCCTTTGCGCTCTTCACTGAGGCAAAATGACTTAAGGTATTAATTCATCT	1440
Db	1381	GCATGACAGCCTTTGCGCTCTTCACTGAGGCAAAATGACTTAAGGTATTAATTCATCT	1440
Qy	1441	ATTTCCTTGGCCACATCTTCTTCTCTGAGCCTACTAATTCATATGCTTATAATTCCAAG	1500
Db	1441	ATTTCCTTGGCCACATCTTCTTCTCTGAGCCTACTAATTCATATGCTTATAATTCCAAG	1500
Qy	1501	CAATGTGCGAAGAAAGAAAGTTAAAGAAATGGAATTAATCCATTTCCCTGTGGCTT	1560
Db	1501	CAATGTGCGAAGAAAGAAAGTTAAAGAAATGGAATTAATCCATTTCCCTGTGGCTT	1560

```
Db 1501 CAATGTCGCAAGAAAGAAATTAAAGATGAAATATTCATTTCCTGTGGCCT 1560
Qy 1561 GTGCGGAGCTGTGTCAGAAACTACTGTCTCCCTTTTCAGACACTCTTGGCCAGAG 1620
Db 1561 GTGCGGAGCTGTGTCAGAAACTACTGTCTCCCTTTTCAGACACTCTTGGCCAGAG 1620
Qy 1621 CAGAGATGAAAAGCCAGGAGGTGAAAGATGATCTTCCAGCTGTGCTCTGTGCG 1680
Db 1621 CAGAGATGAAAAGCCAGGAGGTGAAAGATGATCTTCCAGCTGTGCTCTGTGCG 1680
Qy 1681 AGCCAACTCTTCACTTTGGGGCCCAAGGGGAACTTTTGGAGAGCGCTTGTCTT 1740
Db 1681 AGCCAACTCTTCACTTTGGGGCCCAAGGGGAACTTTTGGAGAGCGCTTGTCTT 1740
Qy 1741 GTACCCCAAGCTGGAATGAGTGGGGGAGTCTCAGCTCACCGCACTCTCCTCTGGG 1800
Db 1741 GTACCCCAAGCTGGAATGAGTGGGGGAGTCTCAGCTCACCGCACTCTCCTCTGGG 1800
Qy 1801 TTCAAGTATTTTCTGCTCAGCTCCCAAGTATGCTGGGAAATACAGGACGCCACATG 1860
Db 1801 TTCAAGTATTTTCTGCTCAGCTCCCAAGTATGCTGGGAAATACAGGACGCCACATG 1860
Qy 1861 CCCAGCTATTTTGTATTTTCAATGAAACGGGATTTTCAACAGCTTGGCCAGCTGTG 1920
Db 1861 CCCAGCTATTTTGTATTTTCAATGAAACGGGATTTTCAACAGCTTGGCCAGCTGTG 1920
Qy 1921 TCGAACTCTGACCCGCAAGTATCCACCGGCTCCGCTCCCAAGATGCTGGGATTAAG 1980
Db 1921 TCGAACTCTGACCCGCAAGTATCCACCGGCTCCGCTCCCAAGATGCTGGGATTAAG 1980
Qy 1981 GGGTGAGCCACCGTGCCTCCGCAAGGAGAACTCTTGTGGAGAGCAGAGGGCTCA 2040
Db 1981 GGGTGAGCCACCGTGCCTCCGCAAGGAGAACTCTTGTGGAGAGCAGAGGGCTCA 2040
Qy 2041 CATCTCCCTCTGATTCCTCCCATGACATTCCTTATCTCTCCCATCTAGCCAGATC 2100
Db 2041 CATCTCCCTCTGATTCCTCCCATGACATTCCTTATCTCTCCCATCTAGCCAGATC 2100
Qy 2101 TATTGTTTCTTCTGTCGCAATTTACTATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 TATTGTTTCTTCTGTCGCAATTTACTATGATGATGATGATGATGATGATGATGAT 2160
Qy 2161 CCCATGGGGGGTGAAGAGGGGTGCAAGGCCCTGCTGCTCACTTTTCTAATTGAA 2220
Db 2161 CCCATGGGGGGTGAAGAGGGGTGCAAGGCCCTGCTGCTCACTTTTCTAATTGAA 2220
Qy 2221 CTGTATTAATATAATCACTTCTGTTTGTTCAGTTTTCAAAAA 2273
Db 2221 CTGTATTAATATAATCACTTCTGTTTGTTCAGTTTTCAAAAA 2273

RESULT 2
AAF25169
ID AAF25169 standard; cDNA; 2264 BP.
XX
AC AAF25169;
XX
DT 30-APR-2001 (first entry)
XX
DE Nucleotide sequence of a human protein having a hydrophobic domain.
XX
KW Human; hydrophobic protein; secretory protein; membrane protein; sepsis;
XX tumour inhibition; immune deficiency; autoimmune disorder; anaemia; burn;
XX infectious disease; cancer; ulcer; periodontal disease; coagulation;
XX Parkinson's disease; fertility; immune response; thrombosis; ss.
XX
OS Homo sapiens.
XX
FH Key 85.1548 Location/Qualifiers
FT CDS
FT /tag= a
FT /product= "hydrophobic protein"
XX
```

```
PN W0200104297-A2.
XX
PD 18-JAN-2001.
XX
PF 16-JUN-2000; 2000WO-JP03942.
XX
PR 08-JUL-1999; 99JP-0194359.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
XX
DR MPI; 2001-103081/11.
XX
P-PSDB; AAB31669.
XX
PT Isolated human proteases and polynucleotides are used in research and
PT have activities including cell proliferation/differentiation activity,
XX immune stimulating activity and receptor/ligand activity -
XX
PS Claim 4; Page 120-124; 151pp; English.
XX
CC The present sequence encodes a human protein with hydrophobic domains.
CC AAF25169 represents a shorter version of the present sequence. The
CC protein possesses a hydrophobic domain and so is a secretory protein
CC or a membrane protein. The protein is used as an antigen to prepare
CC antibodies. The polynucleotide sequence is useful as a source of probes
CC for genetic diagnosis. It is also useful for producing the protein
CC in large quantities and for gene therapy. The eukaryotic cells are used
CC for detecting the receptors or ligands corresponding to the protein and
CC for detecting small novel pharmaceuticals. The antibodies are also used
CC for detection, quantification and purification of the proteins. Both the
CC protein and polynucleotide may be used in research or as nutritional
CC sources or supplements. The protein may have cytokine and cell
CC proliferation/differentiation activity, immune stimulating or suppressing
CC activity, hematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity and tumour inhibition activity. It may therefore be used to
CC treat immune deficiencies resulting from autoimmune disorders or
CC infectious diseases, cancer, sepsis, anaemias, burns and ulcers,
CC periodontal disease, Parkinson's disease, induce fertility, improve
CC immune response and enhance coagulation or inhibit thrombosis.
XX
SQ Sequence 2264 BP; 483 A; 629 C; 550 G; 603 T; 0 other;

Query Match 97.3%; Score 2211; DB 22; Length 2264;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 GGGGGGTCCTGTGGGCTCCCGAGTTAAGATGCGCTCTTCAAGGGGAGAGAGG 106
Db 54 GGGGGGTCCTGTGGGCTCCCGAGTTAAGATGCGCTCTTCAAGGGGAGAGAGG 113
Qy 107 GACTGTGGGCGCTGCGGGGATCTGCAAGCGGGTTCAGAGAGCTGAGCTTAACA 166
Db 114 GACTGTGGGCGCTGCGGGGATCTGCAAGCGGGTTCAGAGAGCTGAGCTTAACA 173
Qy 167 GTTGCAGAGTCCCTGAGCGGCTGAGAACAGGCGTGGCTGATCATCTTCATTTCT 226
Db 174 GTTGCAGAGTCCCTGAGCGGCTGAGAACAGGCGTGGCTGATCATCTTCATTTCT 233
Qy 227 GGGTTAACCCTTGTGTTTATTCGGCATTAATCTTTTCAAGAGAGACTACTACTAT 286
Db 234 GGGTTAACCCTTGTGTTTATTCGGCATTAATCTTTTCAAGAGAGACTACTACTAT 293
Qy 287 CCACCTTTCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGAAACAGCT 346
Db 294 CCACCTTTCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGAAACAGCT 353
Qy 347 CTACCACTCCCTGCTGTGATTTGCTTCACTTCTTCACTTCTTCACTTAATGAGCGCAG 406
Db 354 CTACCACTCCCTGCTGTGATTTGCTTCACTTCTTCACTTCTTCACTTAATGAGCGCAG 413
```

QY 407 CATCATGCCGCTCTCACTACCTTTTGTCTCCAGAGTGGCTTACCTTCTGCTGGATATCA 466
| | | | |
Db 414 CATCATGCCGCTCTCACTACCTTTTGTCTCCAGATGGCTTACCTTCTGCTGGATATCA 473
| | | | |
QY 467 TTACATGCGACCGGCAACTAGATATCAAGTGGAGCAATGCCAATGTGTCTGACTTT 526
| | | | |
Db 474 TTACATGCGACCGGCAACTAGATATCAAGTGGAGCAATGCCAATGTGTCTGACTTT 533
| | | | |
QY 527 GAAGCTGATTTGTTTGGCTGTGCTACTTTTGAACGAGGAAAGATCAGATTTCTTTTC 586
| | | | |
Db 534 GAAGCTGATTTGTTTGGCTGTGCTACTTTTGAACGAGGAAAGATCAGATTTCTTTTC 593
| | | | |
QY 587 CTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCTCTGCTGGAAGTGTGTTT 646
| | | | |
Db 594 CTCTGAGCAACAGAAATATGCCATACGTGTGTTCTTCTCTGCTGGAAGTGTGTTT 653
| | | | |
QY 647 CTCTACTTCTATGAGGCTTCTTGTGAGGCCCCAGTTCTCAATGATCACTACATGAA 706
| | | | |
Db 654 CTCTACTTCTATGAGGCTTCTTGTGAGGCCCCAGTTCTCAATGATCACTACATGAA 713
| | | | |
QY 707 GCTGTGTCAGAGGAGCTGATTTGACATACAGGAAAGATCAAAACAGATCAATTCCTGC 766
| | | | |
Db 714 GCTGTGTCAGAGGAGCTGATTTGACATACAGGAAAGATCAAAACAGATCAATTCCTGC 773
| | | | |
QY 767 TCTCAAGGCTCTGAGTCTGAGCTTTTCTTCTACTAGTGGCTACACACTGCTCAGCCCCA 826
| | | | |
Db 774 TCTCAAGGCTCTGAGTCTGAGCTTTTCTTCTACTAGTGGCTACACACTGCTCAGCCCCA 833
| | | | |
QY 827 CATCAAGAGATCTATCTCTCACTGAAGATATGACAAACCCCTTCTGTTGGCTG 886
| | | | |
Db 834 CATCAAGAGATCTATCTCTCACTGAAGATATGACAAACCCCTTCTGTTGGCTG 893
| | | | |
QY 887 CATGTACATGCTGATCTGAGGCAAGTTTGTCTGTACAAATATGTCACTTGTGGCTGT 946
| | | | |
Db 894 CATGTACATGCTGATCTGAGGCAAGTTTGTGTGTACAAATATGTCACTTGTGGCTGT 953
| | | | |
QY 947 CACAGAAAGATATGATTTTGAAGGCTGAGGCTTCAATGCTTTGAAGAAAGGCA 1006
| | | | |
Db 954 CACAGAAAGATATGATTTTGAAGGCTGAGGCTTCAATGCTTTGAAGAAAGGCA 1013
| | | | |
QY 1007 GGCAGAGTGGAGTGCCTGTGCAACATGAAGTGTGCTCTTGAACAAACCCGCTT 1066
| | | | |
Db 1014 GGCAGAGTGGAGTGCCTGTGCAACATGAAGTGTGCTCTTGAACAAACCCGCTT 1073
| | | | |
QY 1067 CACTGGACCAATGCTCTCAATCAACATCAACCAACGCTGGGTGGCCGCTACATCTT 1126
| | | | |
Db 1074 CACTGGACCAATGCTCTCAATCAACATCAACCAACGCTGGGTGGCCGCTACATCTT 1133
| | | | |
QY 1127 CAAACGATCAAGTCTCTTGGAAATTAAGAACTCTTCAAGGTCTCTGTTGCTATTCCT 1186
| | | | |
Db 1134 CAAACGATCAAGTCTCTTGGAAATTAAGAACTCTTCAAGGTCTCTGTTGCTATTCCT 1193
| | | | |
QY 1187 GGCCCTGTGGCAAGGCTGCACTCAGATACCTGCTGTGCTTCAGATGGAATTCCTCAT 1246
| | | | |
Db 1194 GGCCCTGTGGCAAGGCTGCACTCAGATACCTGCTGTGCTTCAGATGGAATTCCTCAT 1253
| | | | |
QY 1247 TGTATTGTGGAAGAGAGGCTGCAAGCTCATTCAGAGAGGCCCAACCTGAGCAAGCT 1306
| | | | |
Db 1254 TGTATTGTGGAAGAGAGGCTGCAAGCTCATTCAGAGAGGCCCAACCTGAGCAAGCT 1313
| | | | |
QY 1307 GGCCGCAATTAATCTCTCTCCAGCCCTTCTACTATTTTGTGTGAAGACATCCATGGCT 1366
| | | | |
Db 1314 GGCCGCAATTAATCTCTCTCCAGCCCTTCTACTATTTTGTGTGAAGACATCCATGGCT 1373
| | | | |
QY 1367 CTTCATGAGTTACTCCATGACTGCTCTGCTCTTCAACGAGGAAATGGCTTAAGGT 1426
| | | | |
Db 1374 CTTCATGAGTTACTCCATGACTGCTCTGCTCTTCAACGAGGAAATGGCTTAAGGT 1433
| | | | |
QY 1427 GTATTAATCCATCTATTTCTTGGCCATCTTCTTCTGAGCTTACTATTCATATTGCC 1486
| | | | |
Db 1434 GTATTAATCCATCTATTTCTTGGCCATCTTCTTCTGAGCTTACTATTCATATTGCC 1493
| | | | |

QY 1487 TTATATTCAAAAGCAATGCTGCCAAGGAAAGAAAGTTAAAGATGGAATATTCAT 1546
| | | | |
Db 1494 TTATATTCAAAAGCAATGCTGCCAAGGAAAGAAAGTTAAAGATGGAATATTCAT 1553
| | | | |
QY 1547 TTCCCTGTGAGCTGTGCGGGAATGCTGTCAGAAACTACTGTCCTTTTTCACAGACT 1606
| | | | |
Db 1554 TTCCCTGTGAGCTGTGCGGGAATGCTGTCAGAAACTACTGTCCTTTTTCACAGACT 1613
| | | | |
QY 1607 CCTTTGCCAGAGCAGAGATGGAAGATGGAAGCCAGGAGGTGGAAGATGATCTTCAACT 1666
| | | | |
Db 1614 CCTTTGCCAGAGCAGAGATGGAAGATGGAAGCCAGGAGGTGGAAGATGATCTTCAACT 1673
| | | | |
QY 1667 GTGCTCTGCTGTCACCAAGTCTTCAATTTGGGCCCCAAGGGAACTTTTGTGAGA 1726
| | | | |
Db 1674 GTGCTCTGCTGTCACCAAGTCTTCAATTTGGGCCCCAAGGGAACTTTTGTGAGA 1733
| | | | |
QY 1727 AGCGTCTTGTGCTTGTACCAAGTGTGATGAGTGGGAGATCTCAGCTCACCCCAAC 1786
| | | | |
Db 1734 AGCGTCTTGTGCTTGTACCAAGTGTGATGAGTGGGAGATCTCAGCTCACCCCAAC 1793
| | | | |
QY 1787 CTCACCTCTGAGGTTCAAGTATTTTCTGCTCAGCTCCCAAGTGTGGAATACA 1846
| | | | |
Db 1794 CTCACCTCTGAGGTTCAAGTATTTTCTGCTCAGCTCCCAAGTGTGGAATACA 1853
| | | | |
QY 1847 GGCAGCCCAACATGCCAGCTATTTTGTATTTAGATGAAACGGGATTTACACACGT 1906
| | | | |
Db 1854 GGCAGCCCAACATGCCAGCTATTTTGTATTTAGATGAAACGGGATTTACACACGT 1913
| | | | |
QY 1907 TGCCAGGCTGTCTGCTGAACTCTGACCCGCAAGTGTATCCACCTCCGCTCCCAAG 1966
| | | | |
Db 1914 TGCCAGGCTGTCTGCTGAACTCTGACCCGCAAGTGTATCCACCTCCGCTCCCAAG 1973
| | | | |
QY 1967 TGCTGGATTAACAGGCTGAGCCACCGTCCGCCCCAAGGGGAAACTTTTGTGGAGG 2026
| | | | |
Db 1974 TGCTGGATTAACAGGCTGAGCCACCGTCCGCCCCAAGGGGAAACTTTTGTGGAGG 2033
| | | | |
QY 2027 AGCAGAGGCTGCTACATCTCCCTGATTTCCCCCAATGCAATGTGCTTATCTCCCA 2086
| | | | |
Db 2034 AGCAGAGGCTGCTACATCTCCCTGATTTCCCCCAATGCAATGTGCTTATCTCCCA 2093
| | | | |
QY 2087 TCTAGCAGGAATCTATTTGTTTTTCTTCTGCAATTTACTATGATTTGTATGTCGG 2146
| | | | |
Db 2094 TCTAGCAGGAATCTATTTGTTTTTCTTCTGCAATTTACTATGATTTGTATGTCGG 2153
| | | | |
QY 2147 CTACACACACCCCCCATGAGGAGGTGAGAGGGTGCAGAGCCCTGCTCTCCACTT 2206
| | | | |
Db 2154 CTACACACACCCCCCATGAGGAGGTGAGAGGGTGCAGAGCCCTGCTCTCCACTT 2213
| | | | |
QY 2207 TTTCTACCTTGAAGTATTAATAATCACTTCTGTTTGTTCAGTTT 2257
| | | | |
Db 2214 TTTCTACCTTGAAGTATTAATAATCACTTCTGTTTGTTCAGTTT 2264
| | | | |
RESULT 3
AAZ50889
ID AAZ50889 standard, cDNA; 2234 BP.
XX
AC AAZ50889;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human receptor-associated protein cDNA from Incyte clone 2906971.
XX
XX Human receptor-associated protein; HRAB; Incyte clone 2906971;
XX cytosolic; immunomodulatory; anti-inflammatory; cardiant; antianemic;
XX antiarteriosclerotic; hepatotropic; antiallergic antirheumatic;
XX antiaesthetic; osteopathic; antiallergic; antidiabetic; dermatological;
XX neuroprotective; diagnosis; treatment; prevention; reproductive disorder;
XX cardiovascular; cell proliferative; autoimmune; inflammatory; allergy;
XX gastrointestinal; atherosclerosis; cirrhosis; leukaemia; cancer; AIDS;
XX arthritis; anaemia; asthma; dermatitis; diabetes; osteoporosis;
XX multiple sclerosis; irritable bowel syndrome; ss.

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 63..1526
FT /tag= a
FT /product= "HRAP"
XX
XX WO200008155-A2.
XX
XX 17-FEB-2000.
XX
XX
XX 06-AUG-1999; 99WO-US17777.
XX
XX 07-AUG-1998; 98US-0160065.
XX 01-SEP-1998; 98US-0098703.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Hillman JL, Yue H, Lai P, Tang YT, Gorgone GA, Guegler KJ;
XX Corley NC, Baughn MR;
XX
XX WPI: 2000-205710/18.
XX P-PSDB: AAY69987.
XX
XX
XX New human receptor-associated proteins (HRAP) useful for the diagnosis,
XX treatment and prevention of cell proliferative, autoimmune,
XX inflammatory, reproductive, cardiovascular, and gastrointestinal
XX disorders -
XX
XX Claim 9; Pages 89-90; 99pp; English.
XX
XX The present sequence is a cDNA encoding human receptor-associated protein
XX (HRAP) from Incyte clone 2906971 obtained from THYMOTOS cDNA library.
XX This sequence is expressed in nervous, gastrointestinal
XX CC and reproductive tissues. HRAP has cytostatic, immunomodulatory,
XX antiinflammatory, cardiac, antiarteriosclerotic, hepatotropic,
XX antithrombotic, antirheumatic, osteopathic, antiallergic, antianaemic,
XX CC antidiabetic, dermatological and neuroprotective
XX activities. The present sequence is useful in the diagnosis, treatment
XX CC and prevention of disorders associated with HRAP expression, especially
XX cell proliferative, autoimmune/inflammatory, reproductive,
XX cardiovascular and gastrointestinal disorders (e.g. atherosclerosis,
XX cirrhosis, leukaemia, cancer, AIDS, arthritis, allergies, anaemia,
XX CC asthma, dermatitis, diabetes, osteoporosis, multiple sclerosis and
XX irritable bowel syndrome).
XX
XX Sequence 2234 BP; 473 A; 626 C; 541 G; 594 T; 0 other;
SQ
Query Match 96.2%; Score 2186; DB 21; Length 2234;
Best local Similarity 100.0%; Pred. No. 0;
Matches 2186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 GGGGGGTCCTGTGGGGCTCCGAGTTAAGATGGGCTCTCAGCGAGGGGAGCAAGG 106
DB |||||
DB 32 GGGGGGTCCTGTGGGGCTCCGAGTTAAGATGGGCTCTCAGCGAGGGGAGCAAGG 91
QY 107 GACTGTGGTGGCGCTGGCGGGGTTCTGACGTGGGTTTCAGAGAGCTGAAGCTTAACA 166
DB |||||
DB 92 GACTGTGGTGGCGCTGGCGGGGTTCTGACGTGGGTTTCAGAGAGCTGAAGCTTAACA 151
QY 167 GTTGGGACGTCCTGGGCGCTGAGAACAGGCGCTGGCGCTGATGATCTCATCTTCT 226
DB |||||
DB 152 GTTGGGACGTCCTGGGCGCTGAGAACAGGCGCTGGCGCTGATGATCTCATCTTCT 211
QY 227 GGGTTACCTTTGCTTTGTTTATGGGATTACCTTTTCTACAGAGAGACTGATCT 286
DB |||||
DB 212 GGGTTACCTTTGCTTTGTTTATGGGATTACCTTTTCTACAGAGAGACTGATCT 271
QY 287 CCACTCTTCATACCTTTACAGGCGCTCAATGCTTATTTTAACTTTGAAACAGCT 346
DB |||||
DB 272 CCACTCTTCATACCTTTACAGGCGCTCAATGCTTATTTTAACTTTGAAACAGCT 331
QY 347 CTACACTCCCTGTGATGTGCTTCACTTCTATCTTCACTAATGAGCGGCAC 406

DB |||||
DB 332 CTACACTCCCTGTGATGTGCTTCACTTCTATCTTCACTAATGAGCGGCAC 391
QY 407 CATCACTGGGCTTCATCACTTTTGTCTCCAGATGGGCTACCTTTGGCTGATPACTA 466
DB |||||
DB 392 CATCACTGGGCTTCATCACTTTTGTCTCCAGATGGGCTACCTTTGGCTGATPACTA 451
QY 467 TTACACTGCCAGGCAACTAGATATCAAGTGACCAATGGCCATTTGTTCTGACTT 526
DB |||||
DB 452 TTACACTGCCAGGCAACTAGATATCAAGTGACCAATGGCCATTTGTTCTGACTT 511
QY 527 GAACTGATGTGGTGGCTGTGACTACTTTGACGAGGAGAAAGATCAGAAATTCCTGTC 586
DB |||||
DB 512 GAACTGATGTGGTGGCTGTGACTACTTTGACGAGGAGAAAGATCAGAAATTCCTGTC 571
QY 587 CTCTGAGCAAGAAATATGCCATACGTGATGTTCTTCCCTGGGGAAGTTGCTGTT 646
DB |||||
DB 572 CTCTGAGCAAGAAATATGCCATACGTGATGTTCTTCCCTGGGGAAGTTGCTGTT 631
QY 647 CTCTGATCTTCTATGGGGCTTCTGTGAGGGCCCGAGTTCTGATGATCACTACATGA 706
DB |||||
DB 632 CTCTGATCTTCTATGGGGCTTCTGTGAGGGCCCGAGTTCTGATGATCACTACATGA 691
QY 707 GCTGTGACGAGAGAGCTGATTTGACATACAGAGAAAGATACCAACAGCATATTCCTGC 766
DB |||||
DB 692 GCTGTGACGAGAGAGCTGATTTGACATACAGAGAAAGATACCAACAGCATATTCCTGC 751
QY 767 TCTCAAGGCGCTGATGCTGGGCTTTTCTAAGTGGGCTACACATGCTCAGCCCA 826
DB |||||
DB 752 TCTCAAGGCGCTGATGCTGGGCTTTTCTAAGTGGGCTACACATGCTCAGCCCA 811
QY 827 CATCAGAGAGCTATCTCTCTCACTGAGACTATGACCAACCCCTTGTGTTCCGCTG 886
DB |||||
DB 812 CATCAGAGAGCTATCTCTCTCACTGAGACTATGAGCAACCCCTTGTGTTCCGCTG 871
QY 887 CATGATCATGCTGATCTGGGGAAGTTGTGTGTAACAATATGTCACCTGTTGCTGCT 946
DB |||||
DB 872 CATGATCATGCTGATCTGGGGAAGTTGTGTGTAACAATATGTCACCTGTTGCTGCT 931
QY 947 CACAGAGAGATGATGATTTTACAGGCGCTGGGCTTCAATGCTTTGAGAAAGGCA 1006
DB |||||
DB 932 CACAGAGAGATGATGATTTTACAGGCGCTGGGCTTCAATGCTTTGAGAAAGGCA 991
QY 1007 GCAAGAGTGGAGTCTGTGGCAACATGAAAGTGGCTTGTGAAACCAACCCCGCTT 1066
DB |||||
DB 992 GCAAGAGTGGAGTCTGTGGCAACATGAAAGTGGCTTGTGAAACCAACCCCGCTT 1051
QY 1067 CACTGGCACCATTTGCTCTATTCAACATCAACCAACGCGCTGGTGGCCGCTACATCTT 1126
DB |||||
DB 1052 CACTGGCACCATTTGCTCTATTCAACATCAACCAACGCGCTGGTGGCCGCTACATCTT 1111
QY 1127 CAAAGACTCAAGTTCTTGGAAATTAAGAACTCTCAGGGTCTCTGTTGCTATTCCT 1186
DB |||||
DB 1112 CAAAGACTCAAGTTCTTGGAAATTAAGAACTCTCAGGGTCTCTGTTGCTATTCCT 1171
QY 1187 GGCCTCTGGGACGGGCTGCACTCAGATATCGTGGTCTTCCAGATGGAATTCCTCAT 1246
DB |||||
DB 1172 GGCCTCTGGGACGGGCTGCACTCAGATATCGTGGTCTTCCAGATGGAATTCCTCAT 1231
QY 1247 TGTATTGGAAGAGCAGGCTGCAAGGCTCATTAAGAGAGAGAGAGAGAGAGAGAG 1306
DB |||||
DB 1232 TGTATTGGAAGAGCAGGCTGCAAGGCTCATTAAGAGAGAGAGAGAGAGAGAGAG 1291
QY 1307 GGCCTGATTAATGCTCTCAGGCTTCTAATTTTGTGCAACAGAGAGAGAGAGAGAG 1366
DB |||||
DB 1292 GGCCTGATTAATGCTCTCAGGCTTCTAATTTTGTGCAACAGAGAGAGAGAGAGAG 1351
QY 1367 CTTCATGGGTTACTCATGAGCTTCTGAGCTCTTCAAGTGGGGAACAAATGCTTAAGGT 1426
DB |||||
DB 1352 CTTCATGGGTTACTCATGAGCTTCTGAGCTCTTCAAGTGGGGAACAAATGCTTAAGGT 1411
QY 1427 GTATAATCATCTATTTCTTGGCCAACTCTTCTGAGCTTCAATGATATTCCT 1486
DB |||||

Db 1412 GTATTAATCATCTATTCTTGGCCACATCTTCTCTAGGCTACTATTATTTGCC 1471
QY 1487 TTATATTACAAAGCAATGGTGGCCAAAGAAAGAAAGTAAGAGATGAAATATTCAT 1546
Db 1472 TTATATTACAAAGCAATGGTGGCCAAAGAAAGAAAGTAAGAGATGAAATATTCAT 1531
QY 1547 TTCCCTGTGTGGCTGTGGCGGAGCTGTGACAAAATCTCTGTCTCTCTTTTCAAGACT 1606
Db 1532 TTCCCTGTGTGGCTGTGGCGGAGCTGTGACAAAATCTCTGTCTCTTTTCAAGACT 1591
QY 1607 CCTTTGGCCAGACAGAGATGGAAGCAAGGAGGAGTGAAGATGATGCTTCCAGCT 1666
Db 1592 CCTTTGGCCAGACAGAGATGGAAGCAAGGAGGAGTGAAGATGATGCTTCCAGCT 1651
QY 1667 GTGCTCTGTGTCGACGCAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTTTTTGAGA 1726
Db 1652 GTGCTCTGTGTCGACGCAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTTTTTGAGA 1711
QY 1727 AGGCGTCTGTGTTGTACCCACGCTGGAATGCAATGCGGGGATCTCAGCTCACCGCAAC 1786
Db 1712 AGGCGTCTGTGTTGTACCCACGCTGGAATGCAATGCGGGGATCTCAGCTCACCGCAAC 1771
QY 1787 CTCACCTCTCGGGTTCAGAGATTTCTGCTCAGCTCAGCTCCCAAGTAGCTGGGAATACA 1846
Db 1772 CTCACCTCTCGGGTTCAGAGATTTCTGCTCAGCTCAGCTCCCAAGTAGCTGGGAATACA 1831
QY 1847 GGCACGCCACCATGCCAGCTAATTTTGTATTTTCAAGTAAAGCGGATTTTCCACAGCT 1906
Db 1832 GGCACGCCACCATGCCAGCTAATTTTGTATTTTCAAGTAAAGCGGATTTTCCACAGCT 1891
QY 1907 TGGCAGGCTGTGTGCAATCTCTGACCCGCAAGATGCCACCCGCTCCGCTCCCAAG 1966
Db 1892 TGGCAGGCTGTGTGCAATCTCTGACCCGCAAGATGCCACCCGCTCCGCTCCCAAG 1951
QY 1967 TGCTGGATTAACAGGCGTGAGCCACGCTGCCCGGCCCAAGGGGAAACTCTTGGGAGG 2026
Db 1952 TGCTGGATTAACAGGCGTGAGCCACGCTGCCCGGCCCAAGGGGAAACTCTTGGGAGG 2011
QY 2027 AGCAGAGGGGCTCAATCTCCCTCTGATTTCCCATGACATTTGCTTCTCTCCCA 2086
Db 2012 AGCAGAGGGGCTCAATCTCCCTCTGATTTCCCATGACATTTGCTTCTCTCCCA 2071
QY 2087 TCTAGCCAGAAATATATTTGTTTCTTCTGCAATTTACTAATTTGCTATGTCGCCG 2146
Db 2072 TCTAGCCAGAAATATATTTGTTTCTTCTGCAATTTACTAATTTGCTATGTCGCCG 2131
QY 2147 CTACACACACCCCCCATGGGGGGTGAGAGGGGCTGCTGCTCCACTT 2206
Db 2132 CTACACACACCCCCCATGGGGGGTGAGAGGGGCTGCTGCTCCACTT 2191
QY 2207 TTTCTACTTGGAACTGTATTAGATA 2232
Db 2192 TTTCTACTTGGAACTGTATTAGATA 2217
RESULT 4
ABK84420
ID ABK84420 standard; cDNA; 1842 BP.
XX
XX ABK84420;
AC
XX
XX 14-AUG-2002 (first entry)
DT
XX
XX Human cDNA differentially expressed in granulocytic cells #991.
DE
XX
XX Human; sex; granulocytic cell; DNA chip; bacterial infection;
KM viral infection; parasitic infection; protozoal infection;
KM fungal infection; sterile inflammatory disease; psoriasis;
KM rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KM cardiac reperfusion injury; renal reperfusion injury; ARDS;
KM adult respiratory distress syndrome; inflammatory bowel disease;
KM Crohn's disease; ulcerative colitis; periodontal disease;
KM granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.
XX PN WO200228999-A2.
XX PD 11-APR-2002.
XX PF 03-OCT-2001; 2001MO-US30821.
XX PR 03-OCT-2000; 2000US-237189P.
XX PA (GENE-) GENE LOGIC INC.
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
DR
XX
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity
PS
XX
XX Claim 1; SEQ ID No 991; 114bp; English.
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation (especially chronic) in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection, and
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 1842 BP; 413 A; 525 C; 413 G; 491 T; 0 other;
Query Match 81.0%; Score 1842; DB 24; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 TACTCATTCACCTCTTCCATCTTACAGGCTCTCATCTGTTATTTTAACTTGA 338
Db 1 TACTCATTCACCTCTTCCATCTTACAGGCTCTCATCTGTTATTTTAACTTGA 60
QY 339 AACAGCTTACACACTCCCTGTGTGATGTGCTTCACTTCATCTTCAGACTATG 398

Db 61 AACAGCTTACCACTCCCTGCTGTGTATGTGCTTCAGTTCCTCACTCCGACATAATG 120
Qy 399 GGGCCGACCATCATCTGCGCTCTCACTACCTTTTGGTTCAGATGGGCTACCTTGCGCT 458
Db 121 GGGCCGACCATCATCTGCGCTCTCACTACCTTTTGGTTCAGATGGGCTACCTTGCGCT 180
Qy 459 GGATATCTATTAACCTGCAACCGGCACTACGATATCAAGTATGCAATGCAATGCTATGTT 518
Db 181 GGATATCTATTAACCTGCAACCGGCACTACGATATCAAGTATGCAATGCAATGCTATGTT 240
Qy 519 CTGACTTTAAGCTGATTTGGTGGCTGTGACTTCTTGAACGAGGAGAAAGATCAGAA 578
Db 241 CTGACTTTAAGCTGATTTGGTGGCTGTGACTTCTTGAACGAGGAGAAAGATCAGAA 300
Qy 579 TCGCTGTCTGTAGAGCAAGAAATATGCAATACGAGTGTCTTCCCTGCTGAGAA 638
Db 301 TCGCTGTCTGTAGAGCAAGAAATATGCAATACGAGTGTCTTCCCTGCTGAGAA 360
Qy 639 GGTGTTTCTCTACTTCTATGAGGCTCTTGTGAGGCGCCCACTTCAATGATCAC 698
Db 361 GGTGTTTCTCTACTTCTATGAGGCTCTTGTGAGGCGCCCACTTCAATGATCAC 420
Qy 699 TACATGAACCTGTGCAAGGAGAGCTGATTTGACATACGAGAAAGATACCAACGACATC 758
Db 421 TACATGAACCTGTGCAAGGAGAGCTGATTTGACATACGAGAAAGATACCAACGACATC 480
Qy 759 ATTCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACCTAGTGGGCTACACATGCTC 818
Db 481 ATTCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACCTAGTGGGCTACACATGCTC 540
Qy 819 AGCCCCCATCATCAAGAAAGCTATCTCTCACTGAAGACTATGACAAACCCCTTCTG 878
Db 541 AGCCCCCATCATCAAGAAAGCTATCTCTCACTGAAGACTATGACAAACCCCTTCTG 600
Qy 879 TTCCTGCTGATGATCTGCTGATCTGGGCAAGTTTGTGCTGATCAAAATATGCTACCT 938
Db 601 TTCCTGCTGATGATCTGCTGATCTGGGCAAGTTTGTGCTGATCAAAATATGCTACCT 660
Qy 939 TGGCTGTCTCAAGAGAGATGATCTTTGACGGGCTGGGCTTCAATGGCTTTGAAGA 998
Db 661 TGGCTGTCTCAAGAGAGATGATCTTTGACGGGCTGGGCTTCAATGGCTTTGAAGA 720
Qy 999 AAGGGCAAGCAAGTGGAGTGTGCTGCAACATGAGAGTGTGCTTTGAAACAAC 1058
Db 721 AAGGGCAAGCAAGTGGAGTGTGCTGCAACATGAGAGTGTGCTTTGAAACAAC 780
Qy 1059 CCCCCTTCACTGCGCACTGCTCTATTCACATCAACCAACGCTGGTGGCGCC 1118
Db 781 CCCCCTTCACTGCGCACTGCTCTATTCACATCAACCAACGCTGGTGGCGCC 840
Qy 1119 TACATCTTCAAGACTCAAGTTCCTGGAATAAAGAACTCTCAGGGTCTCTCGTTG 1178
Db 841 TACATCTTCAAGACTCAAGTTCCTGGAATAAAGAACTCTCAGGGTCTCTCGTTG 900
Qy 1179 CTATTCCTGCGCTCTGCGCACTGCACTCAGATATCTGCTGCTTCCAGATGAA 1238
Db 901 CTATTCCTGCGCTCTGCGCACTGCACTCAGATATCTGCTGCTTCCAGATGAA 960
Qy 1239 TTCTCTATTTGATTTGTGAAAGACAGGCTGCCAGCTCAATTCAGAGAGCCCACTG 1298
Db 961 TTCTCTATTTGATTTGTGAAAGACAGGCTGCCAGCTCAATTCAGAGAGCCCACTG 1020
Qy 1299 AGCAAGCTGGCGGCACTTCTGCTCAGGCTTCTACATTTTGGTGAACAGACATC 1358
Db 1021 AGCAAGCTGGCGGCACTTCTGCTCAGGCTTCTACATTTTGGTGAACAGACATC 1080
Qy 1359 CACTGCTCTTCAATGGGTTACTCATGATGCTTCTGCTCTTCACTGCGGCAAAATG 1418
Db 1081 CACTGCTCTTCAATGGGTTACTCATGATGCTTCTGCTCTTCACTGCGGCAAAATG 1140
Qy 1419 CTTAAGGTATTAATCACTATTTCTTGGCCACATCTTCTCTGAGCTACTATTC 1478
Db 1141 CTTAAGGTATTAATCACTATTTCTTGGCCACATCTTCTCTGAGCTACTATTC 1200

Qy 1479 ATATTGCTTATATTCACAAAGCAATGGTGGCCAAAGAAAGAACTTAAGAGATGGA 1538
Db 1201 ATATTGCTTATATTCACAAAGCAATGGTGGCCAAAGAAAGAACTTAAGAGATGGA 1260
Qy 1539 TAATCATTTTCCCTGGTGGCTGTGCGGAGCTGTGCAAGAACTAATGCTCTCTTTC 1598
Db 1261 TAATCATTTTCCCTGGTGGCTGTGCGGAGCTGTGCAAGAACTAATGCTCTCTTTC 1320
Qy 1599 ACAGCACTCTTTTGGCCAGAGCAAGAAATGAAAGCCAGGAGCTGAAATGATGATC 1658
Db 1321 ACAGCACTCTTTTGGCCAGAGCAAGAAATGAAAGCCAGGAGCTGAAATGATGATC 1380
Qy 1659 TTCAGAGTGTGCTGTGCGCAGCAAGTCTTATTTGGGGCCAAAGGGGAACTTTT 1718
Db 1381 TTCAGAGTGTGCTGTGCGCAGCAAGTCTTATTTGGGGCCAAAGGGGAACTTTT 1440
Qy 1719 TTTGAGAAAGGCGCTCTTGTTCACCCAGCTGAAATGCAATGGGAGCTCAGCTC 1778
Db 1441 TTTGAGAAAGGCGCTCTTGTTCACCCAGCTGAAATGCAATGGGAGCTCAGCTC 1500
Qy 1779 ACCGCACTCTCACTCTGCGGTTCAAGTATTTTCTGCTCACTCCCAAGTACTG 1838
Db 1501 ACCGCACTCTCACTCTGCGGTTCAAGTATTTTCTGCTCACTCCCAAGTACTG 1560
Qy 1839 GGAATACAGGACGCGCACATGCGGCTAATTTTGTATTTTCAATGAAAGGGATTT 1898
Db 1561 GGAATACAGGACGCGCACATGCGGCTAATTTTGTATTTTCAATGAAAGGGATTT 1620
Qy 1899 CACCACTTTGGCAGAGCTGTCTGCACTCTGACCGCAAGTATCAACCGCTCCGCG 1958
Db 1621 CACCACTTTGGCAGAGCTGTCTGCACTCTGACCGCAAGTATCAACCGCTCCGCG 1680
Qy 1959 TCCCAAGTGTGGGATTTACAGGCTGAGCCACGCTGCCCGCCAAAGGGGAACTCTT 2018
Db 1681 TCCCAAGTGTGGGATTTACAGGCTGAGCCACGCTGCCCGCCAAAGGGGAACTCTT 1740
Qy 2019 GTGGAGGAGCAGAGGGCTCACTCTCCCTGATTTCCCGCATGCAATTTGCTTATC 2078
Db 1741 GTGGAGGAGCAGAGGGCTCACTCTCCCTGATTTCCCGCATGCAATTTGCTTATC 1800
Qy 2079 TCTCCCATCTAGCCAGAAATCTATTTGTTTCTTCTGCC 2120
Db 1801 TCTCCCATCTAGCCAGAAATCTATTTGTTTCTTCTGCC 1842

RESULT 5
ABN96918
ID ABN96918 standard; DNA; 1842 BP.
XX
AC ABN96918;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3416 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumour; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
XX (GENE-) GENE LOGIC INC.
PA
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI, 2002-426119/45.
XX
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
XX
PS Claim 1; SEQ ID NO 3416; 298bp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumor in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatocytic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WPI
XX at fwp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1842 BP; 413 A; 525 C; 413 G; 491 T; 0 other;
XX
XX Query Match 81.0%; Score 1842; DB 24; Length 1842;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 279 TACCTATCCACCTCTTCCATACCTTTCAGGCTCTCAATTGTTAATTTCATTGGA 338
DB 1 TACCTATCCACCTCTTCCATACCTTTCAGGCTCTCAATTGTTAATTTCATTGGA 60
QY 339 AACAGCTTACACCTCCCTGCTGTATGTGTAGTTCCTCACTTCCTTCATG 398
DB 61 AACAGCTTACACCTCCCTGCTGTATGTGTAGTTCCTCACTTCCTTCATG 120
QY 399 GGGCGACATCACTGCGCTCTCTACTACTCTTTCAGATGCGCTTCTGCT 458
DB 121 GGGCGACATCACTGCGCTCTCTACTACTCTTTCAGATGCGCTTCTGCT 180
QY 459 GATATCTTATCACTGCGACCGGCACTAGCATATCAAGTGGACATGCCATTGTGT 518
DB 181 GATATCTTATCACTGCGACCGGCACTAGCATATCAAGTGGACATGCCATTGTGT 240
QY 519 CTGACTTGAAGCTGATGTTGGCTGTGTGACTTTCAGCGAGGGAAGATCAGAT 578
DB 241 CTGACTTGAAGCTGATGTTGGCTGTGTGACTTTCAGCGAGGGAAGATCAGAT 300
QY 579 TCCCTTCTCTGAGCAAGAAATATGCAATAGTGTCTTCCCTGCTGGAAGT 638
DB 301 TCCCTTCTCTGAGCAAGAAATATGCAATAGTGTCTTCCCTGCTGGAAGT 360
QY 639 GCTGCTTCTCTACTTCTATGGGCTTCTGTTAGGGCCCACTTCTCATGATC 698
DB 361 GCTGCTTCTCTACTTCTATGGGCTTCTGTTAGGGCCCACTTCTCATGATC 420
QY 699 TACATGAAGCTGGGAGGAGCTATGATACACGGAAGATACCAAGATC 758
DB 421 TACATGAAGCTGGGAGGAGCTATGATACACGGAAGATACCAAGATC 480
QY 759 ATTCTCTCTCAAGCGCTGAGTCTGGGCTTTTCTTACCTAGTGGCTACACTGCTC 818
DB 481 ATTCTCTCTCAAGCGCTGAGTCTGGGCTTTTCTTACCTAGTGGCTACACTGCTC 540
QY 819 AGCCCAATCAAGAACTATCTCTCACTGAAGATATGCAACACCCCTTCTGG 878
DB 541 AGCCCAATCAAGAACTATCTCTCACTGAAGATATGCAACACCCCTTCTGG 600
QY 879 TTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 938

DB 601 TTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 939 TGGCTGTGACAGAAAGATATGATTTTTCAGGCTGCTGCTTCAATGCTTGAAGA 998
DB 661 TGGCTGTGACAGAAAGATATGATTTTTCAGGCTGCTGCTTCAATGCTTGAAGA 720
QY 999 AAGGCAAGGCAAGTGGGATGCTGTGCCAATGATGATGATGATGATGATGATGAT 1058
DB 721 AAGGCAAGGCAAGTGGGATGCTGTGCCAATGATGATGATGATGATGATGATGAT 780
QY 1059 CCGGCTTACAGGCAAGTGGGATGCTGTGCCAATGATGATGATGATGATGATGATGAT 1118
DB 781 CCGGCTTACAGGCAAGTGGGATGCTGTGCCAATGATGATGATGATGATGATGATGAT 840
QY 1119 TACATCTTCAAGACTCAAGTCTTGGAAATTAAGAACTCTCTGAGGCTCTCTGTT 1178
DB 841 TACATCTTCAAGACTCAAGTCTTGGAAATTAAGAACTCTCTGAGGCTCTCTGTT 900
QY 1179 CTATTTCTGCTCTGCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1238
DB 901 CTATTTCTGCTCTGCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 1239 TTCCTATTGTTATGTTGGAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
DB 961 TTCCTATTGTTATGTTGGAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1299 AGCAAGCTGCGCCCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
DB 1021 AGCAAGCTGCGCCCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1359 CATGCTCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1418
DB 1081 CATGCTCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1419 CTTAAGGTGTAATATCAATCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1478
DB 1141 CTTAAGGTGTAATATCAATCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1479 ATAATGCTTATATTAATCAAGAAAGATGCTGCAAGAAAGATTAAGAAAGATGAA 1538
DB 1201 ATAATGCTTATATTAATCAAGAAAGATGCTGCAAGAAAGATTAAGAAAGATGAA 1260
QY 1539 TAATCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1598
DB 1261 TAATCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1599 ACAGCACTCTTGGCCCAAGAGAGAAATGAAAGCCAGGAGGTGGAAGATGATGATG 1658
DB 1321 ACAGCACTCTTGGCCCAAGAGAGAAATGAAAGCCAGGAGGTGGAAGATGATGATG 1380
QY 1659 TTCAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1718
DB 1381 TTCAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1719 TTTGGAAGAGGCTTGTGCTTGTGCAACGCTGGAATGCAATGCGGGAATCTCAGCTC 1778
DB 1441 TTTGGAAGAGGCTTGTGCTTGTGCAACGCTGGAATGCAATGCGGGAATCTCAGCTC 1500
QY 1779 ACCGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1838
DB 1501 ACCGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY 1839 GGAATACAGGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1898
DB 1561 GGAATACAGGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
QY 1899 CACCAAGTGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1958
DB 1621 CACCAAGTGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1959 TCCCAAGTGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2018

Db 1681 TCCCAAGTCTGGATTACAGCGTGAGCACCGTCCGCGCCCAAGGGGAACTCTT 1740
QY GTGGAGAGAGAGAGGGGCTCACATCTCCCTCTGATTCCTCCATGACATTGCTTATC 2078
Db 1741 GTGGAGAGAGAGAGGGGCTCACATCTCCCTCTGATTCCTCCATGACATTGCTTATC 1800
QY 2079 TCTCCCATCTAGCAGGAATCTATTGTGTTTCTTCTGCCC 2120
Db 1801 TCTCCCATCTAGCAGGAATCTATTGTGTTTCTTCTGCCC 1842

RESULT 6
ABL64098
ID ABL64098 standard; DNA; 1842 BP.
XX
AC ABL64098;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:2435.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinos;
KW cytoelastic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX

05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235638P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237588P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.

PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX MPI, 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 2435; 44bp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (II)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytoelastic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1842 BP; 413 A; 525 C; 413 G; 491 T; 0 other;

Query Match 81.0%; Score 1842; DB 24; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 TACCTCATCACCTCTTCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTGA 338
Db 1 TACCTCATCACCTCTTCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTGA 60

QY 339 AACCACTTACCACTCCTCTGCTGATTTGCTTCACTTCTTCACTTCTTGA 398
Db 61 AACCACTTACCACTCCTCTGCTGATTTGCTTCACTTCTTCACTTCTTGA 120

QY 399 GGGCGACCATACAGCGCTCTCTCACTTCTTTCAGATGCGCTTCTGCGT 458
Db 121 GGGCGACCATACAGCGCTCTCTCACTTCTTTCAGATGCGCTTCTGCGT 180

QY 459 GGATACCTATTACAGTCCACCGCAACTGATATCAAGTGACATGCCAATTGT 518
Db 181 GGATACCTATTACAGTCCACCGCAACTGATATCAAGTGACATGCCAATTGT 240

QY 519 CTGACTTTGAAGCTGATTTGCTTGGCTTGACTTCTTGAAGGAGAAATCAGAT 578
Db 241 CTGACTTTGAAGCTGATTTGCTTGGCTTGACTTCTTGAAGGAGAAATCAGAT 300

QY 579 TCGTGTCTCTGAGGAAGAAATATGCAATACGAGGCTCTTCCCTGCTGAAAT 638
Db 301 TCGTGTCTCTGAGGAAGAAATATGCAATACGAGGCTCTTCCCTGCTGAAAT 360

QY 639 GCTGTTTCTCTTACTTATGAGGCTCTTCTGTAAGGCCAGTTCTCAATGATC 698
Db 361 GCTGTTTCTCTTACTTATGAGGCTCTTCTGTAAGGCCAGTTCTCAATGATC 420

QY 699 TACATGAAGCTGTGAGGAGAGCTGATGACATACGAAAGATACCAACGATC 758
Db 421 TACATGAAGCTGTGAGGAGAGCTGATGACATACGAAAGATACCAACGATC 480

QY 759 ATTCTGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACTAGTGGGCTACACATGCTC 818
 DB 481 ATTCCGCTCTCAAGGCGCTGAGTCTGGGCTTTTCTACTAGTGGGCTACACATGCTC 540
 QY 819 AGCCCCACATCAAGAGACTATCTCTCACTGAAAGTATGACAAACCCCTCTGG 878
 DB 541 AGCCCCACATCAAGAGACTATCTCTCACTGAAAGTATGACAAACCCCTCTGG 600
 QY 879 TTCGCTGATGATGATCTGATCTGGGCAAGTTTGTCTGTACAAATATGTCACTGT 938
 DB 601 TTCGCTGATGATGATCTGATCTGGGCAAGTTTGTCTGTACAAATATGTCACTGT 660
 QY 939 TGGCTGTGACAGAGAGATGATGATTTTGAACGGGCTTGGGCTTCAATGGCTTTGAAGA 998
 DB 661 TGGCTGTGACAGAGAGATGATGATTTTGAACGGGCTTGGGCTTCAATGGCTTTGAAGA 720
 QY 999 AAGGCAAGGCAAGTGGAGTGGCTGTGCCAATGAGGTGGCTTTTGAACAAAC 1058
 DB 721 AAGGCAAGGCAAGTGGAGTGGCTGTGCCAATGAGGTGGCTTTTGAACAAAC 780
 QY 1059 CCCCCTTCACTGGACCAATTCCTCATTCACATCAACCAACGCTGGGCTGG 1118
 DB 781 CCCCCTTCACTGGACCAATTCCTCATTCACATCAACCAACGCTGGGCTGG 840
 QY 1119 TACATCTTCAAGAGACTCAAGTCTCTTGAATAAAGAACTCTCTCAAGGTCTCTGTTG 1178
 DB 841 TACATCTTCAAGAGACTCAAGTCTCTTGAATAAAGAACTCTCTCAAGGTCTCTGTTG 900
 QY 1179 CTATTCCTGGCCCTCTGGACGCGCTGCACTCAAGATACCTGATCTCTTCCAGATGGA 1238
 DB 901 CTATTCCTGGCCCTCTGGACGCGCTGCACTCAAGATACCTGATCTCTTCCAGATGGA 960
 QY 1239 TTCCTCATTTGATTTGTGGAAGACAGGCTGCCAGGCTCATTAAGAGAGCCCTG 1298
 DB 961 TTCCTCATTTGATTTGTGGAAGACAGGCTGCCAGGCTCATTAAGAGAGCCCTG 1020
 QY 1299 AGCAAGCTGGCGGCTTACTGCTCCAGCCCTTCTACTATTTGGTGGACAGACATC 1358
 DB 1021 AGCAAGCTGGCGGCTTACTGCTCCAGCCCTTCTACTATTTGGTGGACAGACATC 1080
 QY 1359 CACTGCTCTTCAATGGGTTACTCATGACTGCTCTGCTCTTCAAGTGGAGCAATGG 1418
 DB 1081 CACTGCTCTTCAATGGGTTACTCATGACTGCTCTGCTCTTCAAGTGGAGCAATGG 1140
 QY 1419 CTTAAGGTGATAAATCATCTATTTCTTGGCCACATCTTCTGAGCTACTATTC 1478
 DB 1141 CTTAAGGTGATAAATCATCTATTTCTTGGCCACATCTTCTGAGCTACTATTC 1200
 QY 1479 ATATTGCTTATTTTCAAGAGATGGTGGCAAGAGAAAGATTAAGAGATGGA 1538
 DB 1201 ATATTGCTTATTTTCAAGAGATGGTGGCAAGAGAAAGATTAAGAGATGGA 1260
 QY 1539 TAATCATTTTCTGCTGGCTGTGGGGAATGTCAGAAATCTGCTCTCTTTTC 1598
 DB 1261 TAATCATTTTCTGCTGGCTGTGGGGAATGTCAGAAATCTGCTCTCTTTTC 1320
 QY 1599 ACAGCACTCTTTTCCCGCAGAGAGATGAAAGCCAGGAGGTGAAGATCATGC 1658
 DB 1321 ACAGCACTCTTTTCCCGCAGAGAGATGAAAGCCAGGAGGTGAAGATCATGC 1380
 QY 1659 TTCCAGCTGTGCTGTGGCCAGCAAGTCTTATTTGGGGCAAGGGGAATCTTTT 1718
 DB 1381 TTCCAGCTGTGCTGTGGCCAGCAAGTCTTATTTGGGGCAAGGGGAATCTTTT 1440
 QY 1719 TTTGAGAGAGCGCTTGTGTTGTACCCAGCTGGAATGATGAGCGGATCTCAGCTC 1778
 DB 1441 TTTGAGAGAGCGCTTGTGTTGTACCCAGCTGGAATGATGAGCGGATCTCAGCTC 1500
 QY 1779 ACCGCAACTCTCACTCTGCTGGTTCAAGATATTTTCTGCTCAGCTCTCCAGATGCTG 1838
 DB 1501 ACCGCAACTCTCACTCTGCTGGTTCAAGATATTTTCTGCTCAGCTCTCCAGATGCTG 1560
 QY 1839 GGAATACAGGCAAGCCACCATGCGCAGCTAATTTTGTATTTCTAGTAAAGCGGATTT 1898

DB 1561 GGAATACAGGCAAGCCACCATGCGCAATATTTTGTATTTTGTATGATAAAGGGATTT 1620
 QY 1899 CACCAAGTTGGCAGAGCTGTGCTGCAACTCTGACCGCAAGTATGATCAACCGCTCCGCC 1958
 DB 1621 CACCAAGTTGGCAGAGCTGTGCTGCAACTCTGACCGCAAGTATGATCAACCGCTCCGCC 1680
 QY 1959 TCCCAAGTTGCTGGATTACAGGCGTGAACCACTGCTCCCGGCCAAGGGGAATCTTT 2018
 DB 1681 TCCCAAGTTGCTGGATTACAGGCGTGAACCACTGCTCCCGGCCAAGGGGAATCTTT 1740
 QY 2019 GTGGAGAGAGAGAGGCGCTCATCTCCCTCTGATTTCCCATGACATTTGCTTATC 2078
 DB 1741 GTGGAGAGAGAGAGGCGCTCATCTCCCTCTGATTTCCCATGACATTTGCTTATC 1800
 QY 2079 TCTCCCATCTAGCCAGCAATCTATTTGTTTCTTCTGCC 2120
 DB 1801 TCTCCCATCTAGCCAGCAATCTATTTGTTTCTTCTGCC 1842

RESULT 7
 AAF25159
 ID AAF25159 standard; cDNA; 1461 BP.
 XX
 AC AAF25159;
 XX
 DT 30-APR-2001 (first entry)
 XX
 XX
 DE Nucleotide sequence of a human protein having a hydrophobic domain.
 XX
 KW Human; hydrophobic protein; secretory protein; membrane protein; sepsis;
 KW tumour inhibition; immune deficiency; autoimmune disorder; anaemia; burn;
 KW infectious disease; cancer; ulcer; peridontal disease; coagulation;
 KW Parkinson's disease; fertility; immune response; thrombosis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1461
 FT CDS /tag= a
 FT /product= "hydrophobic protein"
 FT /note= "no termination codon given"
 XX
 XX
 PN WO200104297-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 16-JUN-2000; 2000MO-JP03942.
 XX
 PR 08-JUL-1999; 99JP-0194359.
 XX
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI; 2001-103081/11.
 DR P-PSDB; AAB31669.
 XX
 PT Isolated human proteins and polynucleotides are used in research and
 PT have activities including cell proliferation/differentiation activity,
 PT immune stimulating activity and receptor/ligand activity -
 XX
 PS Claim 3; Page 108-109; 151pp; English.
 XX
 CC The present sequence encodes a human protein with hydrophobic domains.
 CC AAF25159 represents a longer version of the present sequence. The
 CC protein possesses a hydrophobic domain and so is a secretory protein
 CC or a membrane protein. The protein is used as an antigen to prepare
 CC antibodies. The polynucleotide sequence is useful as a source of probes
 CC for genetic diagnosis. It is also useful for producing the protein
 CC in large quantities and for gene therapy. The eukaryotic cells are used
 CC for detecting the receptors or ligands corresponding to the protein and

CC for detecting small novel pharmaceuticals. The antibodies are also used
 CC for detection, quantification and purification of the proteins. Both the
 CC protein and polynucleotide may be used in research or as nutritional
 CC sources or supplements. The protein may have cytokine and cell
 CC proliferation/differentiation activity, immune stimulating or suppressing
 CC activity, hematopoiesis regulating activity, tissue growth or activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity and tumour inhibition activity. It may therefore be used to
 CC treat immune deficiencies resulting from autoimmune disorders or
 CC infectious diseases, cancer, sepsis, anaemias, burns and ulcers,
 CC periodontal disease, Parkinson's disease, induce fertility, improve
 CC immune response and enhance coagulation or inhibit thrombosis.

Sequence 1461 BP; 326 A; 401 C; 336 G; 398 T; 0 other;

Query Match 64.3%; Score 1461; DB 22; Length 1461;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 ATGGGCTCTCAGCGGAGGGGAGACGAGGAGCTGTGTGGCGCTGGCGGGGTTCTGCG 137
 Db 1 ATGGGCTCTCAGCGGAGGGGAGACGAGGAGCTGTGTGGCGCTGGCGGGGTTCTGCG 60
 QY 138 TCGGGTTTCCAGAGCTGAGCCTTAACAAGTTGGCGACGTCCCTGGGCGCGTCAGAACAG 197
 Db 61 TCGGGTTTCCAGAGCTGAGCCTTAACAAGTTGGCGACGTCCCTGGGCGCGTCAGAACAG 120
 QY 198 GCGCTGCGGCTATCATCTTCATCTTCTGTGGGTTACCCCTTTGTTTATGGCAT 257
 Db 121 GCGCTGCGGCTATCATCTTCATCTTCTGTGGGTTACCCCTTTGTTTATGGCAT 180
 QY 258 TACCTTTTCTAAGAGACCTACCTCATCACTCCCTTTCATACCTTTACAGCCCTCTCA 317
 Db 181 TACCTTTTCTAAGAGACCTACCTCATCACTCCCTTTCATACCTTTACAGCCCTCTCA 240
 QY 318 ATTGCTATTATTTAATCTTGAACACGACTCTACCACTCCCTGTGTATGTGCTTCA 377
 Db 241 ATTGCTATTATTTAATCTTGAACACGACTCTACCACTCCCTGTGTATGTGCTTCA 300
 QY 378 TTCTCTATCTTGTGACTAATGGGCGGACCATCATCTGCGGCTCTCACTATTTGCTTC 437
 Db 301 TTCTCTATCTTGTGACTAATGGGCGGACCATCATCTGCGGCTCTCACTATTTGCTTC 360
 QY 438 CAGATGGCTTACCTTCTGGCTGATATCTATTAACCTGACACCGGCAACATGATATCAAG 497
 Db 361 CAGATGGCTTACCTTCTGGCTGATATCTATTAACCTGACACCGGCAACATGATATCAAG 420
 QY 498 TCGACATATCCACATTTGTCTGACTTTGAAGCTGATGTTGGCTGTGACTACTTT 557
 Db 421 TCGACATATCCACATTTGTCTGACTTTGAAGCTGATGTTGGCTGTGACTACTTT 480
 QY 558 GACGAGGAGAAAGATCAGAAATTCCTTGTCTGTAGCAACAGAAATATGCAATACGTGT 617
 Db 481 GACGAGGAGAAAGATCAGAAATTCCTTGTCTGTAGCAACAGAAATATGCAATACGTGT 540
 QY 618 GTTCTTCCCTGCTGGAAGTGTGTTTCTCTCACTTTTAAGGGCCCTTTTGTGAAGG 677
 Db 541 GTTCTTCCCTGCTGGAAGTGTGTTTCTCTCACTTTTAAGGGCCCTTTTGTGAAGG 600
 QY 678 CCCGATTTCTCATGATCATCTACTAATGAAGCTGTGACGAGAGCTGATTGACATCA 737
 Db 601 CCCGATTTCTCATGATCATCTACTAATGAAGCTGTGACGAGAGCTGATTGACATCA 660
 QY 738 GGAAGATATCCAAACAGCATCTTCTGTCTCAAGCGCTGAGTGTGGGCTTTTCTAC 797
 Db 661 GGAAGATATCCAAACAGCATCTTCTGTCTCAAGCGCTGAGTGTGGGCTTTTCTAC 720
 QY 798 CTAGTGGGTACACACTGTGACGCGCCCATCAACAGAGATATCTCTCACTGAAGC 857
 Db 721 CTAGTGGGTACACACTGTGACGCGCCCATCAACAGAGATATCTCTCACTGAAGC 780
 QY 858 TATGACAAACACCCCTTGTGTGCTGCTCATGTATCATCTGTGAGGCAAGTTTGG 917

Db 781 TATGACAAACACCCCTTGTGTGCTGCTCATGTATCATGTGATCTGGGCAAGTTTGG 840
 QY 918 CTGTCAAAATATGTACCTGTTGGCTGTGTACAGAGAGATGTGATTTTGAACGGGCTG 977
 Db 841 CTGTCAAAATATGTACCTGTTGGCTGTGTGTACAGAGAGATGTGATTTTGAACGGGCTG 900
 QY 978 GGCCTCAATGTGCTTGAAGAAAGGCAAGGCAAGTGGAGTGGCTGTGCGCAACATGAAG 1037
 Db 901 GGCCTCAATGTGCTTGAAGAAAGGCAAGGCAAGTGGAGTGGCTGTGCGCAACATGAAG 960
 QY 1038 GTGTGGCTCTTTGAAGAAACAAACCCCGCTTCACTGTGACACCATTTCTTCAACATCA 1097
 Db 961 GTGTGGCTCTTTGAAGAAACAAACCCCGCTTCACTGTGACACCATTTCTTCAACATCA 1020
 QY 1098 ACCAAGCCTGTGGTGGCGCGCTACATCTTCAACAGCATCTCAAGTCTTGGAAATGAAG 1157
 Db 1021 ACCAAGCCTGTGGTGGCGCGCTACATCTTCAACAGCATCTCAAGTCTTGGAAATGAAG 1080
 QY 1158 CTCTCTCAAGGCTCTGTGTGTATCTGTGGCGCTGTGGACAGGCTGACTCAGATAC 1217
 Db 1081 CTCTCTCAAGGCTCTGTGTGTATCTGTGGCGCTGTGGACAGGCTGACTCAGATAC 1140
 QY 1218 CTGTCTGTCTTCAAGATGAAATTCCTCATTTGTTATTTGAAAGACAGGCTGCCAGCTC 1277
 Db 1141 CTGTCTGTCTTCAAGATGAAATTCCTCATTTGTTATTTGAAAGACAGGCTGCCAGCTC 1200
 QY 1278 ATTCAAGAGAGCCCACTTGTGACAGAGCTGCGGCAATTAATCTGTCTTCAAGCCTTTC 1337
 Db 1201 ATTCAAGAGAGCCCACTTGTGACAGAGCTGCGGCAATTAATCTGTCTTCAAGCCTTTC 1260
 QY 1338 TATTGTGCAACACATCCAGGCTTTCATGAGGTTACTGCATGACGCTTTCG 1397
 Db 1261 TATTGTGCAACACATCCAGGCTTTCATGAGGTTACTGCATGACGCTTTCG 1320
 QY 1398 CTCTTCAAGTGGCAAAATGCTTAAGTGTATTAATCATATTTCTTGGCCACATC 1457
 Db 1321 CTCTTCAAGTGGCAAAATGCTTAAGTGTATTAATCATATTTCTTGGCCACATC 1380
 QY 1458 TTCTCTGAGCTTCAATTCATATTTGCTTATTTACAAAGCAATGTGCAAGGAA 1517
 Db 1381 TTCTCTGAGCTTCAATTCATATTTGCTTATTTACAAAGCAATGTGCAAGGAA 1440
 QY 1518 GAGAGTTAAGAGATGAA 1538
 Db 1441 GAGAGTTAAGAGATGAA 1461

RESULT 8
 AAS70385
 ID AAS70385 standard; cDNA, 1867 BP.
 XX
 AC AAS70385;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #6189.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 PN W0200175067-A2.
 FN
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.

DB 1561 GCCTCCCAAGTAGCTGGAAATACAGSCACGCCATGCCAGCTATTTTGTATTTTC 1620
QY 1883 AATTAGAAAGGATTTTACCAGCTTGGCCAGAGCTGTCTCGAATCTCTTACCGCAGTA 1942
DB 1621 AATTAGAAAGGATTTTACCAGCTTGGCCAGAGCTGTCTCGAATCTCTTACCGCAGTA 1680
QY 1943 TCCACCGCGCTCGCTCCCAAGTCTGGATTTACAGGCGTG 1985
DB 1681 TCCACCGCGCTCGCTCCCAAGTCTGGATTTACAGGCGTG 1723

RESULT 9
AAS30639/C
ID AAS30639 standard; DNA; 7461 BP.
XX AAS30639;
XX 21-NOV-2001 (first entry)
XX
DE DNA encoding novel lung cancer antigen, Seq ID No 91.
XX
XX Human; lung cancer; immunosuppressive; antiarthritic; antirheumatic;
XX antiproliferative; cytostatic; cardiatic; vasotropic; cerebroprotective;
XX neuroprotective; antibacterial; virucide; fungicide;
XX ophthalmological; autoimmune disease; rheumatoid arthritis; neoplasm;
XX hyperproliferative disorder; cardiovascular disorder; angiodenesis;
XX nervous system disorder; Alzheimer's disease; infection; skin aging;
XX ocular disorder; wound healing; organ transplantation; ds.
XX
XX Homo sapiens.
XX
XX MO200155300-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01238.
XX
XX 31-JAN-2000; 2000US-0179065.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX MPI; 2001-465565/50.
XX
XX Isolated nucleic acid molecule encoding a lung cancer antigen is used
XX in preventing, treating or ameliorating a medical condition
XX
XX Disclosure: SEQ ID No 91; 475bp; English.
XX
XX The invention relates to novel isolated lung cancer antigen
XX polynucleotide (I) and polypeptides (II). (I) and (II) are used to
XX prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX rabbits, goats, cats, dogs, chickens or sheep. (I) and (II) are
XX also used in diagnosing a pathological condition or susceptibility to a
XX pathological condition, in particular lung cancer. The antibodies to
XX (II) can also be used in alleviating symptoms associated with the
XX disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
XX linked immunosorbent assays (ELISA). Disorders which are diagnosed or
XX treated include autoimmune diseases e.g. rheumatoid arthritis,
XX hyperproliferative disorders e.g. neoplasms of the breast or liver,
XX cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
XX e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi and
XX ocular disorders e.g. corneal infection. The polypeptides can also be
XX used to aid wound healing and epithelial cell proliferation, to prevent
XX skin aging due to sunburn, to maintain organs before transplantation, for
XX supporting cell culture of primary tissues, to regenerate tissues and in
XX chemotaxis. AAS30580-AAS30655 represent novel human lung cancer antigen
XX coding sequences, PCR primers and related sequences of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly

CC from WIFO at: ftp.wifo.int/pub/published_pct_sequences.
XX
SQ Sequence 7461 BP; 1883 A; 1666 C; 1923 G; 1989 T; 0 other:
Query Match 31.1%; Score 708; DB 22; Length 7461;
Best Local Similarity 100.0%; Pred. No. 3e-300;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GGTGGCTGTGGGGGACGTGGGAGAACTCTGCTCCCTTTTACAGCACTCTTGG 1612
DB 714 GGTGGCTGTGGGGGACGTGGGAGAACTCTGCTCCCTTTTACAGCACTCTTGG 655
QY 1613 CCCAGACAGAGAAATGAAAAAGCCAGAGAGGTGAGATCGATCTTCCAGCTGTGCT 1672
DB 654 CCCAGACAGAGAAATGAAAAAGCCAGAGAGGTGAGATCGATCTTCCAGCTGTGCT 595
QY 1673 CTGCTGCCAGCCCACTTTTCAATTTGGGGCCAAAGGGGAACTTTTGGAGAAAGCGT 1732
DB 594 CTGCTGCCAGCCCACTTTTCAATTTGGGGCCAAAGGGGAACTTTTGGAGAAAGCGT 535
QY 1733 CTGCTTTGTATCACCACGCTGGAAATGACAGTGGCGGGATCTCAGCTCACCGCACTCCAC 1792
DB 534 CTGCTTTGTATCACCACGCTGGAAATGACAGTGGCGGGATCTCAGCTCACCGCACTCCAC 475
QY 1793 CTGCTGGGTTCAAGGATTTTCTGCTCAGGCTCCCAAGTGTGGGAAATACAGGACAG 1852
DB 474 CTGCTGGGTTCAAGGATTTTCTGCTCAGGCTCCCAAGTGTGGGAAATACAGGACAG 415
QY 1853 CCACCATGCCCACTAAATTTTGTATTTTCAATGAGAAACGGGATTTACACAGTTGGCCA 1912
DB 414 CCACCATGCCCACTAAATTTTGTATTTTCAATGAGAAACGGGATTTACACAGTTGGCCA 355
QY 1913 GGTGCTGTCCAACTCTCGACCGGCAAGTATCACCCGCTCCGCTCCCAAGTGTGG 1972
DB 354 GGTGCTGTCCAACTCTCGACCGGCAAGTATCACCCGCTCCGCTCCCAAGTGTGG 295
QY 1973 GATTACAGGCGTGACCCACCGTCCCGGCCCAAGGGGAACTCTTGTGGAGAGACAGA 2032
DB 294 GATTACAGGCGTGACCCACCGTCCCGGCCCAAGGGGAACTCTTGTGGAGAGACAGA 235
QY 2033 GGGGCTCACTCTCCCTCTGATTTCCCATGACATTTGCTTCTCCCATCTAGC 2092
DB 234 GGGGCTCACTCTCCCTCTGATTTCCCATGACATTTGCTTCTCCCATCTAGC 175
QY 2093 CAGGAATCATTTGTTTCTTCTGCAATTTATATGATTTGATGCGGCTACCA 2152
DB 174 CAGGAATCATTTGTTTCTTCTGCAATTTATGATTTGATGCGGCTACCA 115
QY 2153 CCACCCCGCCCATGGGGGGGTGAGAGGGGTGCAAGCCCTGCTCCATCTTTCTA 2212
DB 114 CCACCCCGCCCATGGGGGGGTGAGAGGGGTGCAAGCCCTGCTCCATCTTTCTA 55
QY 2213 CTTTGAACCTGATTAGATTAATATCATCTTGTGTTGATGTTTCA 2260
DB 54 CTTTGAACCTGATTAGATTAATATCATCTTGTGTTGATGTTTCA 7

RESULT 10
AAS28701/C
ID AAS28701 standard; DNA; 7461 BP.
XX AAS28701;
XX 07-NOV-2001 (first entry)
XX
DE Genomic sequence #541 encoding for novel human respiratory antigen.
XX
XX Human; respiratory antigen; respiratory disorder; throat disorder;
XX lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX respiratory active; ds.
XX
XX Homo sapiens.
XX

XX MO200155448-A1.
 PD 02-AUG-2001.
 XX 17-JAN-2001; 2001WO-US01333.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 XX

PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241121.
 PR 20-OCT-2000; 2000US-0241185.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 06-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX

21-SEP-2000; 2000US-224274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 13-OCT-2000; 2000US-239935P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 PI MPI; 2003-311001/30.
 XX
 PT Novel lung cancer related polypeptide useful for detecting, treating,
 PT preventing and/or prognosing lung, respiratory, hyperproliferative,
 PT immune system, cardiovascular endocrine or gastrointestinal disorders
 XX
 PS Disclosure; Page 282-286; 332pp; English.
 XX
 CC The invention relates to an isolated lung cancer related polypeptide its
 CC variants, allelic variants or homologues. The polypeptide and its nucleic
 CC acid are useful for preventing, treating, or ameliorating a medical
 CC condition in a mammalian subject, for diagnosing a pathological condition
 CC or a susceptibility to a pathological condition in a subject by
 CC determining the presence or absence of mutation in the nucleic acid or
 CC biological sample and/or amount of expression of the polypeptide in a
 CC result. The polypeptide, antibodies to the polypeptide or the
 CC polynucleotide are useful for detecting, treating, preventing and/or
 CC prognosing disorders of the lung, such as small cell lung cancer, non-
 CC small cell lung cancer (e.g. squamous cell carcinoma, adenocarcinoma,
 CC large cell carcinoma, adenocarcinoma and undifferentiated
 CC carcinoma) or lung cancer metastasis. The polypeptide, its antibodies or
 CC its polynucleotide are also useful for detecting, treating, preventing
 CC and/or prognosing respiratory disorders e.g. nonallergic rhinitis and
 CC adult respiratory distress syndrome; hyperproliferative disorders e.g.
 CC lung neoplasms and prostate neoplasms; immune system disorders e.g.
 CC rheumatoid arthritis and multiple sclerosis; blood-related disorders e.g.
 CC anaemia and leukopenia; allergic reactions e.g. anaphylaxis and asthma;
 CC inflammatory disorders e.g. appendicitis and dermatitis; graft vs host
 CC disease; immune complex disease e.g. serum sickness and polyarteritis
 CC nodosa; urinary system disorders e.g. glomerulonephritis and kidney
 CC failure; cardiovascular disorders e.g. arrhythmia and myocardial
 CC infarction; musculoskeletal system disorders e.g. Albers-Schönberg
 CC disease and osteoarthritis; neurological disorders e.g. Alzheimer's
 CC disease and Parkinson's disease; endocrine disorders e.g. Addison's
 CC disease and diabetes mellitus; gastrointestinal disorders e.g. ulcers and
 CC gastric reflux; reproductive system disorders e.g. infertility and
 CC impotence; developmental and inherited disorders e.g. Fanconi's syndrome
 CC and cerebral palsy; diseases at the cellular level e.g. cancer and
 CC amyotrophic lateral sclerosis; infectious diseases e.g. viral and
 CC bacterial. The polypeptide, its antibodies or its polynucleotide are also
 CC useful for wound healing e.g. grafts and burns and for regeneration of
 CC tissues damaged due to e.g. aging and trauma. The present sequence

CC represents DNA encoding a human lung cancer antigen.
 XX
 SQ Sequence 7461 BP, 1883 A, 1666 C, 1923 G, 1989 T; 0 other:
 Query Match 31.1%; Score 708; DB 25; Length 7461;
 Best Local Similarity 100.0%; Pred. No. 3e-300;
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1553 GGTGGCTTGGGGGAGTGTGAGAACTACTCTGCTTCCCTTTTACAGCACTCTCTTGG 1612
 DB 714 GGTGGCTTGGGGGAGTGTGAGAACTACTCTGCTTCCCTTTTACAGCACTCTCTTGG 655
 QY 1613 CCCGAGAGAGAAAGAAAGCCAGGGAGGAGAGATGATGCTTCAGCTGAGCT 1672
 DB 654 CCCGAGAGAGAAAGAAAGCCAGGGAGGAGAGATGATGCTTCAGCTGAGCT 595
 QY 1673 CTGCTCCAGCCAGCAAGCTTCTATTTGGGGCCAAAGGGGAACTTTTGGAGAGCGCT 1732
 DB 594 CTGCTCCAGCCAGCAAGCTTCTATTTGGGGCCAAAGGGGAACTTTTGGAGAGCGCT 535
 QY 1733 CTGCTTTTGTACCCAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 1792
 DB 534 CTGCTTTTGTACCCAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 475
 QY 1793 CTGCTGGGTTCAGATATTTCTGCTCAGCTCCCAAGTATGATGATGATGATGATGATG 1852
 DB 474 CTGCTGGGTTCAGATATTTCTGCTCAGCTCCCAAGTATGATGATGATGATGATGATG 415
 QY 1853 CCACATGCCAGCAATTTTGTATTTTCTAGTAGAAGGAGATTTTACACAGCTTGAGCA 1912
 DB 414 CCACATGCCAGCAATTTTGTATTTTCTAGTAGAAGGAGATTTTACACAGCTTGAGCA 355
 QY 1913 GGTGCTTCTGAACTCTCTGACCCGAGATGATGATGATGATGATGATGATGATGATG 1972
 DB 354 GGTGCTTCTGAACTCTCTGACCCGAGATGATGATGATGATGATGATGATGATGATG 295
 QY 1973 GATTACAGGGGTGAGGACCGGTCCCGGCCAAGAGGAACTCTTGTGGAGAGAGAGA 2032
 DB 294 GATTACAGGGGTGAGGACCGGTCCCGGCCAAGAGGAACTCTTGTGGAGAGAGAGA 235
 QY 2033 GGGGCTCACATCTCCCTCTGATTCGCCCATGACATGCTTATCTCTCCCATTTAGC 2092
 DB 234 GGGGCTCACATCTCCCTCTGATTCGCCCATGACATGCTTATCTCTCCCATTTAGC 175
 QY 2093 CAGGAATCTATTTGTTTTTCTTCTGCAATTTACTATGATTTGTATGTCGGCTACCA 2152
 DB 174 CAGGAATCTATTTGTTTTTCTTCTGCAATTTACTATGATTTGTATGTCGGCTACCA 115
 QY 2153 CACACCCCGCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTGCTCCACTTTTCTA 2212
 DB 114 CACACCCCGCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTGCTCCACTTTTCTA 55
 QY 2213 CTTTGAAGTATTTAGATTAATCACTTCTGTTTTCAGTTTCA 2260
 DB 54 CTTTGAAGTATTTAGATTAATCACTTCTGTTTTCAGTTTCA 7
 RESULT 12
 ABA1565/c
 ID ABA1565 standard; DNA; 32174 BP.
 AC ABA1565;
 XX
 DT 23-JUN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 7996.
 XX
 KW Human, noctropic; neuroprotective; cytoskeletal; dermatological; vitruide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antiscidling; antianaemic; antiallergic; cancer;
 KW antitubercular; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;

KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX Homo sapiens.
XX
XX WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251859.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX
PS Disclosure; SEQ ID NO 7996; 1701bp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 32174 BP; 8001 A; 6837 C; 7064 G; 8747 T; 1525 other;
Query Match 31.1%; Score 708; DB 22; Length 32174;
Best Local Similarity 100.0%; Pred. No. 2.8e-300;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GGTGGCTGTGGCGGAGTGTGTCAGAAACTACTGCTCTCCCTTTTCACAGACTCTTTG 1612
DB 6043 GGTGGCTGTGGCGGAGTGTGTCAGAAACTACTGCTCTCCCTTTTCACAGACTCTTTG 5984
QY 1613 CCCAGAGCAGAGATGGAAGAGCAGAGAGTGAAGTGAAGTTCACAGCTGTGCT 1672
DB 5983 CCCAGAGCAGAGATGGAAGAGCAGAGAGTGAAGTGAAGTTCACAGCTGTGCT 5924
QY 1673 CTGCTGCACCAAGTCTTCACTTTGGGCAAGAGGAACTTTTGGAGAGGCGT 1732
DB 5923 CTGCTGCACCAAGTCTTCACTTTGGGCAAGAGGAACTTTTGGAGAGGCGT 5864
QY 1733 CTGCTGTGTACCCAGCTGGAATGCAAGTGGCGGATCTCACTCAACCGCACTTCAC 1792
DB 5863 CTGCTGTGTACCCAGCTGGAATGCAAGTGGCGGATCTCACTCAACCGCACTTCAC 5804
QY 1793 CTCTGGGTTCAAGTATTTTCTGCTCAAGCCCTCCCAAGAGTGGGAATACAGGAGC 1852
DB 5803 CTCTGGGTTCAAGTATTTTCTGCTCAAGCCCTCCCAAGAGTGGGAATACAGGAGC 5744
QY 1853 CCAACATGCCAGCTAATTTTGTATTTTCAAGTGAAGAGGATTTCCACAGTTGGCCA 1912
DB 5743 CCAACATGCCAGCTAATTTTGTATTTTCAAGTGAAGAGGATTTCCACAGTTGGCCA 5684
QY 1913 GGTGGTGTGAACTCTGACCCGCAAGTATCCACCGGCTCGCTCCCAAGTGTGG 1972
DB 5683 GGTGGTGTGAACTCTGACCCGCAAGTATCCACCGGCTCGCTCCCAAGTGTGG 5624
QY 1973 GATTACAGGGGTAGCAGCGTGCAGGCGCCAAAGGGGAAACTTTGTGGAGAGAGAGA 2032
DB 5623 GATTACAGGGGTAGCAGCGTGCAGGCGCCAAAGGGGAAACTTTGTGGAGAGAGAGA 5564
QY 2033 GGGGCTACATCTCTCTGATTTCCCAATGACATTTGCTTCTCCCATTTAGC 2092
DB 5563 GGGGCTACATCTCTCTGATTTCCCAATGACATTTGCTTCTCCCATTTAGC 5504

QY 2093 CAGGAATCTATTGTGTTTCTTCTTCGCCAATTACTATGATGTGATGCGGCTACCA 2152
DB 5503 CAGGAATCTATTGTGTTTCTTCTTCGCCAATTACTATGATGTGATGCGGCTACCA 5444
QY 2153 CCACCCCCCCCAAGGGGGGTGAGAGGGGTGCAAGGCCCTGCTCCACTTTTCTA 2212
DB 5443 CCACCCCCCCCAAGGGGGGTGAGAGGGGTGCAAGGCCCTGCTCCACTTTTCTA 5384
QY 2213 CCTTGAAGTGTATTTAGATAAATCACTTCTGTTGTGACGTTTTC 2260
DB 5383 CCTTGAAGTGTATTTAGATAAATCACTTCTGTTGTGACGTTTTC 5336
RESULT 13
ABAI9477
ID ABAI9477 standard; DNA; 32174 BP.
XX ABAI9477;
AC
XX
XX 23-JAN-2002 (first entry)
DT
XX
DE Human nervous system related polynucleotide SEQ ID NO 11808.
XX
KW Human; nootropic; neuroprotective; cyrostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischizoid; antianemic; antitubercular; cancer;
KW antipneumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotoxic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN MO200159063-A2.
XX
PD 16-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225457.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.

22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234224.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 01-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249208.

17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Disclosure; SEQ ID NO 11808; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 32174 BP; 8747 A; 7064 C; 6837 G; 8001 T; 1525 other;

Query Match 31.1%; Score 708; DB 22; Length 32174;
 Best Local Similarity 100.0%; Pred. No. 2.8e-300;
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GGTGGCTGTGGGAGCTGTGCAGAAATCTACTGCTCCCTTTTACAGCACTCTTGG 1612
 Db 26132 GGTGGCTGTGGGAGCTGTGCAGAAATCTACTGCTCCCTTTTACAGCACTCTTGG 26191
 QY 1613 CCCAGAGCAGAGAAATGAGAAAGCCAGGAGGTGAGATGATGATCTTCCAGCTGTGCT 1672

```
Db 26192 CCCAGAGCAGAAATGAAAGCAGGAGGTGAAGATGATGCTTCCAGCTGTGCTT 26251
Qy 1673 CTCTGCGCAGCAAGTCTTCATTTGGGCCCAAGGGGAACTTTTGGAGAGGCGT 1732
Db 26252 CTCTGCGCAGCAAGTCTTCATTTGGGCCCAAGGGGAACTTTTGGAGAGGCGT 26311
Qy 1733 CTTCCTTTGTCAACCGCTGGAATGAGTGGGGATCTCAGCTCACCGCACTTCAC 1792
Db 26312 CTTCCTTTGTCAACCGCTGGAATGAGTGGGGATCTCAGCTCACCGCACTTCAC 26371
Qy 1793 CTCTGGGTCAGATGATTTTCGCTCAGCCTCCCAAGTAGTGGAAATACAGGACG 1852
Db 26372 CTCTGGGTCAGATGATTTTCGCTCAGCCTCCCAAGTAGTGGAAATACAGGACG 26431
Qy 1853 CCACATGCCAGCTAATTTTGTATTTTCAGTAAGAAACGGGATTTCCACACGTTGCGCA 1912
Db 26432 CCACATGCCAGCTAATTTTGTATTTTCAGTAAGAAACGGGATTTCCACACGTTGCGCA 26491
Qy 1913 GGCTGGTCTGAACTCTGACCGCAATGATCCACCGCTCCGCTCCCAAGTGTGG 1972
Db 26492 GGCTGGTCTGAACTCTGACCGCAATGATCCACCGCTCCGCTCCCAAGTGTGG 26551
Qy 1973 GATTACAGGGGTAGCAGCAGTGCAGGCGCCCAAGGGGAAACTCTTGGAGAGGACAGA 2032
Db 26552 GATTACAGGGGTAGCAGCAGTGCAGGCGCCCAAGGGGAAACTCTTGGAGAGGACAGA 26611
Qy 2033 GGGGCTCAGATCTCCCTCTGATTCGCCCATGACATGCTTATCTCCCATCTAGC 2092
Db 26612 GGGGCTCAGATCTCCCTCTGATTCGCCCATGACATGCTTATCTCCCATCTAGC 26671
Qy 2093 CAGGAATCTATTTGTTTCTTCTGCAATTATCTATGATGTAATGTCGCTACCA 2152
Db 26672 CAGGAATCTATTTGTTTCTTCTGCAATTATCTATGATGTAATGTCGCTACCA 26731
Qy 2153 CCAACCCCCCATGGGGGGGAGAGGGGTCAGAGCCCTGCTCCACTTTTCTA 2212
Db 26732 CCAACCCCCCATGGGGGGGAGAGGGGTCAGAGCCCTGCTCCACTTTTCTA 26791
Qy 2213 CCTTGAATCTATTAATAAATCACTTCGTTGTTCAAGTTTCA 2260
Db 26792 CCTTGAATCTATTAATAAATCACTTCGTTGTTCAAGTTTCA 26839

RESULT 14
ABA20359 standard; DNA; 32174 BP.
XX ABA20359;
AC ABA20359;
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 12690.
XX
KW Human; noctropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; antitubercular; anti-HIV; antibacterial; antiviral;
KW antiparkinsonian; antileukemic; antianemic; antileukemic; cancer;
KW antitubercular; hepatotropic; cerebroprotective; antineurotoxic;
KW antileukemic; antidiabetic; antileukemic; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotoxic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234273.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
```


RESULT 15
ABA21505/c
ID ABA21505 standard; DNA; 32174 BP.
XX
AC ABA21505;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 13836.
XX
KW Human; nootropic; neuroprotective; cyostatic; dermatological; viroicide;
KW immunosuppressive; anti-HIV; antibacterial; vulnerrary;
KW antiparkinsonian; antischlingic; antianaemic; antidiabetic; cancer;
KW antithrombotic; hepatocytic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilicer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231247.
PR 08-SEP-2000; 2000US-0231247.
PR 08-SEP-2000; 2000US-0231247.

PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.

THIS PAGE BLANK (USPTO)

Db 61 GGGCTCCGGAGTTAAGATGGCGTCTCAGCGGAGGGGGACGAGGGGACTGTGTGGCCG 120
Qy 121 TGGCGGGGGGTTCTGCACTCGGGTTTCCAGAGCTGAACCTTAAACAAGTGGCGACGTCC 180
Db 121 TGGCGGGGGGTTCTGCACTCGGGTTTCCAGAGCTGAACCTTAAACAAGTGGCGACGTCC 180
Qy 181 TGGGGCGGTGAGAACGAGGGGCTGGGGTATCATCTCCATCTTCGGGGTTACCCCTTG 240
Db 181 TGGGGCGGTGAGAACGAGGGGCTGGGGTATCATCTTCATCTTCGGGGTTACCCCTTG 240
Qy 241 CTGTGTTTATCGGCATTAACCTTTTCTACAGAGAGACCTACCTACCTCTTCCATA 300
Db 241 CTGTGTTTATCGGCATTAACCTTTTCTACAGAGAGACCTACCTACCTCTTCCATA 300
Qy 301 CCTTTACAGGCTCTCAATTGCTTATTTAATCTTTGAAACAGCTCTACACCTCCCTGC 360
Db 301 CCTTTACAGGCTCTCAATTGCTTATTTAATCTTTGAAACAGCTCTACACCTCCCTGC 360
Qy 361 TGTGTAATGTGCTTCACTTCTCATCTCTGACCTAAAGGGGCGGACCATCATCTGGCTGC 420
Db 361 TGTGTAATGTGCTTCACTTCTCATCTCTGACCTAAAGGGGCGGACCATCATCTGGCTGC 420
Qy 421 TCACATCTTTTCTCTTCAGATGGCTTCTGCTGGATATCTATTACACTGCCACG 480
Db 421 TCACATCTTTTCTCTTCAGATGGCTTCTGCTGGATATCTATTACACTGCCACG 480
Qy 481 GCAACTACGATTCAGATGGAACAATGCACTGTGTCTGACTTTGAACCTGATGGTT 540
Db 481 GCAACTACGATTCAGATGGAACAATGCACTGTGTCTGACTTTGAACCTGATGGTT 540
Qy 541 TGGCTGTGCTACTTTGACGAGGAGGAAATGAGATTCCTTCTCTGACCAACAGA 600
Db 541 TGGCTGTGCTACTTTGACGAGGAGGAAATGAGATTCCTTCTCTGACCAACAGA 600
Qy 601 AATATGCAATACGTGTGTCTCTCCCTGCTGAGATGCTGTGTTCTCTACTTCTATG 660
Db 601 AATATGCAATACGTGTGTCTCTCCCTGCTGAGATGCTGTGTTCTCTACTTCTATG 660
Qy 661 GGGGCTTTCTGTGAGGGGCGGAGTTCTCAATGAATCACTACATGAGCTGTGAGGAG 720
Db 661 GGGGCTTTCTGTGAGGGGCGGAGTTCTCAATGAATCACTACATGAGCTGTGAGGAG 720
Qy 721 AGCTGATGCAATACGAGAAATACCAACAGCATCTCTGCTCTCAAGCGCTGA 780
Db 721 AGCTGATGCAATACGAGAAATACCAACAGCATCTCTGCTCTCAAGCGCTGA 780
Qy 781 GTCTGGGCTTTTCTACTAGTGGGCTACACACTGTCAAGCCCATCATCAGAGAGCT 840
Db 781 GTCTGGGCTTTTCTACTAGTGGGCTACACACTGTCAAGCCCATCATCAGAGAGCT 840
Qy 841 ATCTCTCACTGAAGATATGACCAACCCCTTCTGTGCTCCGCTGATATGATGCTGA 900
Db 841 ATCTCTCACTGAAGATATGACCAACCCCTTCTGTGCTCCGCTGATATGATGCTGA 900
Qy 901 TCTGGGGCAAGTTTGTGCTGACAAATATGACCTGTGCTGTGCTGACAGAGAGAT 960
Db 901 TCTGGGGCAAGTTTGTGCTGACAAATATGACCTGTGCTGTGCTGACAGAGAGAT 960
Qy 961 GCATTTTGAAGGGGCTTGAATGCTTTTGAAGAAAGGCGCAAGGTGGAGT 1020
Db 961 GCATTTTGAAGGGGCTTGAATGCTTTTGAAGAAAGGCGCAAGGTGGAGT 1020
Qy 1021 CTTGTGCAACATGAGAGGTGTGCTTTTGAACAAACCCCGCTTCACTGGGACCATG 1080
Db 1021 CTTGTGCAACATGAGAGGTGTGCTTTTGAACAAACCCCGCTTCACTGGGACCATG 1080
Qy 1081 CCTCATTCACATCAACCAAGCGCTGGGTGGCGGCTACATCTTAAAGCATCAAGT 1140
Db 1081 CCTCATTCACATCAACCAAGCGCTGGGTGGCGGCTACATCTTAAAGCATCAAGT 1140
Qy 1141 TCCCTTGAATTAAGAACTCTCTCAGGGGTCTCTGTGTATTTCTGGCCCTCTGGGACG 1200
Db 1141 TCCCTTGAATTAAGAACTCTCTCAGGGGTCTCTGTGTATTTCTGGCCCTCTGGGACG 1200

Qy 1201 GCTTGCACTCAGGATACCTGTGTCTTCCAGATGGAATTCCTCATTTGTTATGTGAAA 1260
Db 1201 GCTTGCACTCAGGATACCTGTGTCTTCCAGATGGAATTCCTCATTTGTTATGTGAAA 1260
Qy 1261 GACAGGCTGCAGGCTCATTTCAAGAGAGCCCAACCTGAGCAAGCTGGCGGCACTACTG 1320
Db 1261 GACAGGCTGCAGGCTCATTTCAAGAGAGCCCAACCTGAGCAAGCTGGCGGCACTACTG 1320
Qy 1321 TCCCTCAGGCTTCTACTATTTGTGTGCAACAGACATCCACTGTGCTTTCAATGGATTACT 1380
Db 1321 TCCCTCAGGCTTCTACTATTTGTGTGCAACAGACATCCACTGTGCTTTCAATGGATTACT 1380
Qy 1381 CCATGACCTGCTTCTGCTTCTTCACTGGGACAAATGCTTTAAGGTGTATTAATCATCT 1440
Db 1381 CCATGACCTGCTTCTGCTTCTTCACTGGGACAAATGCTTTAAGGTGTATTAATCATCT 1440
Qy 1441 ATTTCTTGGGCAATCTTCTCTGAGGCTATTCATTTGCTTATTTCACAAG 1500
Db 1441 ATTTCTTGGGCAATCTTCTCTGAGGCTATTCATTTGCTTATTTCACAAG 1500
Qy 1501 CAATGTGCAAGAGAAAGAAATTAAAGAGATGGAATTAATTCATTTCTGTGGCT 1560
Db 1501 CAATGTGCAAGAGAAAGAAATTAAAGAGATGGAATTAATTCATTTCTGTGGCT 1560
Qy 1561 GTGCGGACCTGTGTGCAAGAACTACTGTCTCTCTTTTCAAGACCTCTTGGCCAGAG 1620
Db 1561 GTGCGGACCTGTGTGCAAGAACTACTGTCTCTCTTTTCAAGACCTCTTGGCCAGAG 1620
Qy 1621 CAGAGAAATGAAAAGGCGAGGAGGTGGAAGATGATTCAGAGCTGTGCTGTGCTGC 1680
Db 1621 CAGAGAAATGAAAAGGCGAGGAGGTGGAAGATGATTCAGAGCTGTGCTGTGCTGC 1680
Qy 1681 AGCAAGCTTCTCATTTTGGGCGCAAGAGGAGAACTTTTGTGAGAGGCGTCTTCT 1740
Db 1681 AGCAAGCTTCTCATTTTGGGCGCAAGAGGAGAACTTTTGTGAGAGGCGTCTTCT 1740
Qy 1741 GTACCCCAAGCTGGAATGCAAGTGGGGATCTCAGCTCACCGCACTCTCTCTGG 1800
Db 1741 GTACCCCAAGCTGGAATGCAAGTGGGGATCTCAGCTCACCGCACTCTCTCTGG 1800
Qy 1801 TTCAGATATTTTCTGCTCAGGCTCCCAAGATGAGTGGGAAATACAGGACGCAACATG 1860
Db 1801 TTCAGATATTTTCTGCTCAGGCTCCCAAGATGAGTGGGAAATACAGGACGCAACATG 1860
Qy 1861 CCCAGCTAATTTTGTATTTTCAAGTAAACGGAATTTTCAACAGCTGTGCGCAGGCTGTC 1920
Db 1861 CCCAGCTAATTTTGTATTTTCAAGTAAACGGAATTTTCAACAGCTGTGCGCAGGCTGTC 1920
Qy 1921 TCGAATCTCTGACCGCAAGTATCCACCGGCTCCGCTCCCAAGATGCTGGGATTAAG 1980
Db 1921 TCGAATCTCTGACCGCAAGTATCCACCGGCTCCGCTCCCAAGATGCTGGGATTAAG 1980
Qy 1981 GCGTAGGACCAAGTGTGCGGCGCCAAAGGAGAACTTTGTGAGAGGAGAGAGGCTCA 2040
Db 1981 GCGTAGGACCAAGTGTGCGGCGCCAAAGGAGAACTTTGTGAGAGGAGAGAGGCTCA 2040
Qy 2041 CATCTCCCTGTGATTCGCCCATGACATTTGCTTCTCCCATGAGCAGGAATC 2100
Db 2041 CATCTCCCTGTGATTCGCCCATGACATTTGCTTCTCCCATGAGCAGGAATC 2100
Qy 2101 TATTTGTTTTTCTTCTGCAATTTATCTATGATTTGTGTGTGCGGCTACCAACACCCC 2160
Db 2101 TATTTGTTTTTCTTCTGCAATTTATCTATGATTTGTGTGTGCGGCTACCAACACCCC 2160
Qy 2161 CCCATGGGGGGGTGAGAGGGGTGCAAGGCGCTGCTCCACTTTTCTACTCTTGGA 2220
Db 2161 CCCATGGGGGGGTGAGAGGGGTGCAAGGCGCTGCTCCACTTTTCTACTCTTGGA 2220
Qy 2221 CTGTATTAATAAATCACTTCTGTGTGTCAAGTTTCAAAAAA 2273
Db 2221 CTGTATTAATAAATCACTTCTGTGTGTCAAGTTTCAAAAAA 2273

RESULT 2
US-09-880-107-3415
; Sequence 3415, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3415
; LENGTH: 1842
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U72515
US-09-880-107-3415

Query Match 81.0%; Score 1842; DB 10; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 TACCCATCCACCTCTTCCATCTTACAGGCTCTCAATGCTTATTTTACCTTTGA 338
DB 1 TACCTATCCACCTCTTCCATCTTACAGGCTCTCAATGCTTATTTTACCTTTGA 60
QY 339 AACGACTTACACCTCTCTGCTGATATGCTTCACTTCTGATCTTCTGATTAAG 398
DB 61 AACGACTTACACCTCTCTGCTGATATGCTTCACTTCTGATCTTCTGATTAAG 120
QY 339 GGGCCACCATACCTGCGCTCTCACTTCTTCTGCTTCTGATGCTTCTGCTGCT 458
DB 121 GGGCCACCATACCTGCGCTCTCACTTCTTCTGCTTCTGATGCTTCTGCTGCT 180
QY 459 GGATCTATTACCTGCGCAGGCACTACATGATGCAATGCAATGCAATGCTTGT 518
DB 181 GGATCTATTACCTGCGCAGGCACTACATGATGCAATGCAATGCAATGCTTGT 240
QY 519 CTGACTTGAAGCTGATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 578
DB 241 CTGACTTGAAGCTGATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 300
QY 579 TCTTGTCTCTGAGCAACAGAAATATGCAATGCTGCTTCTTCTGCTGGAAGTT 638
DB 301 TCTTGTCTCTGAGCAACAGAAATATGCAATGCTGCTTCTTCTGCTGGAAGTT 360
QY 639 GGTGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 698
DB 361 GGTGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420
QY 699 TACATGAAGCTGCTGAGGAGAGCTGATGCTTGTGCTTGTGCTTGTGCTTGTGCT 758
DB 421 TACATGAAGCTGCTGAGGAGAGCTGATGCTTGTGCTTGTGCTTGTGCTTGTGCT 480
QY 759 ATTCTGCTCTCAAGGCTGAGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 818
DB 481 ATTCTGCTCTCAAGGCTGAGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
QY 819 AGCCCAATCATCAAGAGACTATCTCTCACTGAAGACTATGACCAACCCCTTCTG 878
DB 541 AGCCCAATCATCAAGAGACTATCTCTCACTGAAGACTATGACCAACCCCTTCTG 600
QY 879 TTCGCTGCTGATGCTGATCTGAGGCAAGTTTGTGCTGATCAAAATATGCTGCTGT 938

DB 601 TTCGCTGCTGATGCTGATCTGAGGCAAGTTTGTGCTGATCAAAATATGCTGCTGT 660
QY 939 TGGCTGCTCAAGAGAGATGCTTGTGCTGCTGCTGCTTCAATGCTTGTGAAGA 998
DB 661 TGGCTGCTCAAGAGAGATGCTTGTGCTGCTGCTGCTTCAATGCTTGTGAAGA 720
QY 999 AAGGCAAGGCAAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
DB 721 AAGGCAAGGCAAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 1059 CCCCCTTCACTGAGCACCATTGCTTCAATCAATCAACCAACGCTGAGGCTGCTGCT 1118
DB 781 CCCCCTTCACTGAGCACCATTGCTTCAATCAATCAACCAACGCTGAGGCTGCTGCT 840
QY 1119 TACATCTTCAAGCACTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1178
DB 841 TACATCTTCAAGCACTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1179 CTATTCCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1238
DB 901 CTATTCCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 1239 TTCTCTATTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
DB 961 TTCTCTATTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1299 AGCAAGCTGCGGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
DB 1021 AGCAAGCTGCGGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1359 CACTGCTCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1418
DB 1081 CACTGCTCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1419 CTAAAGTGTATTAATCCATCTATTTCTTGGCCACATCTTCTGCTGCTGCTGCTGCT 1478
DB 1141 CTAAAGTGTATTAATCCATCTATTTCTTGGCCACATCTTCTGCTGCTGCTGCTGCT 1200
QY 1479 ATATTGCTTATTTATTCACAAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1538
DB 1201 ATATTGCTTATTTATTCACAAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1539 TAATCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1598
DB 1261 TAATCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1599 ACAGCACTCTTTTGGCCAGAGAGAGATGGAAGCAAGGCTGCTGCTGCTGCTGCTGCT 1658
DB 1321 ACAGCACTCTTTTGGCCAGAGAGAGATGGAAGCAAGGCTGCTGCTGCTGCTGCTGCT 1380
QY 1659 TTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1718
DB 1381 TTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1719 TTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1778
DB 1441 TTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1779 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1838
DB 1501 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
QY 1839 GGAATACAGGCAAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1898
DB 1561 GGAATACAGGCAAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1899 CACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1958
DB 1621 CACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
QY 1959 TTCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2018

Db 1681 TCCCAAGTCTGGATTAACAGGCGTGGACCGCTGCCGCCCAAGGGAACACTT 1740
Qy 2019 GTGGAGAGACAGAGGGGCTCACATCTCCCTCTGATTTCCCATGACATTGCTTATC 2078
Db 1741 GTGGAGAGACAGAGGGGCTCACATCTCCCTCTGATTTCCCATGACATTGCTTATC 1800
Qy 2079 TCTCCCATCTACAGCAAGATCTATGTGTCTTCTTCTGSCC 2120
Db 1801 TCTCCCATCTACAGCAAGATCTATGTGTCTTCTTCTGSCC 1842

RESULT 3
US-09-954-531-1368
Sequence 1368, Application US/09954531
Patent No. US20020165180A1
GENERAL INFORMATION:
APPLICANT: Weaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
TITLE OF INVENTION: Gene Sets
FILE REFERENCE: 689290-77
CURRENT APPLICATION NUMBER: US/09/954,531
CURRENT FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1368
LENGTH: 1842
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-1368

Query Match 81.0%; Score 1842; DB 10; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 TACCTCATCACTCTTCCATACCTTTACAGGCTCTCAATTGCTTATTTAACTTTGA 338
Db 1 TACCTCATCACTCTTCCATACCTTTACAGGCTCTCAATTGCTTATTTAACTTTGA 60
Qy 339 AACCAAGCTACCACTCCGCTGTGTATTTGCTTCAAGTTCCTCATCTTCACTATG 398
Db 61 AACCAAGCTACCACTCCGCTGTGTATTTGCTTCAAGTTCCTCATCTTCACTATG 120
Qy 399 GGGCGACATCACTGCGCTCTCTCACTACCTTTTGTCCAGATGCGCTACCTTGTGCT 458
Db 121 GGGCGACATCACTGCGCTCTCTCACTACCTTTTGTCCAGATGCGCTACCTTGTGCT 180
Qy 459 GGATCACTTACATGCGACCGGCACTAGATTAAGTGGAGAAATGCCACATTTGTT 518
Db 181 GGATCACTTACATGCGACCGGCACTAGATTAAGTGGAGAAATGCCACATTTGTT 240
Qy 519 CTGACTTTGAAGCTGATTTGTTGGCTGTGACTACTTTTGAAGGAGGAAAGATCAGAT 578
Db 241 CTGACTTTGAAGCTGATTTGTTGGCTGTGACTACTTTTGAAGGAGGAAAGATCAGAT 300
Qy 579 TCTTGTCTCTGAGCAAGAAATATGCAATAGTGTGTTCTTCCCTGCTGGAAGTT 638
Db 301 TCTTGTCTCTGAGCAAGAAATATGCAATAGTGTGTTCTTCCCTGCTGGAAGTT 360
Qy 639 GCTGTGTTCTCTACTTCTATGAGGCTTCTTGTGAGGCGCCAGTTCTCAATAATCAC 698
Db 361 GCTGTGTTCTCTACTTCTATGAGGCTTCTTGTGAGGCGCCAGTTCTCAATAATCAC 420
Qy 699 TACATGAAGCTGTGAGGAGAGCTGATTTGACATACAGGAAGATACCAACAGCATC 758

Db 421 TACATGAAGCTGTGAGGAGAGAGCTGATTTGACATACAGGAAGATACCAACAGCATC 480
Qy 759 ATTCTGCTCTCAAGGCGCTGAGTGTGGGCTTTTCTACCTAGTGGCTACACATGCTC 818
Db 481 ATTCTGCTCTCAAGGCGCTGAGTGTGGGCTTTTCTACCTAGTGGCTACACATGCTC 540
Qy 819 AGCCCCCAATACAGAAAGATCTATCTCTCACTGAAAGACTATAGAACCAACCCCTTCTGG 878
Db 541 AGCCCCCAATACAGAAAGATCTATCTCTCACTGAAAGACTATAGAACCAACCCCTTCTGG 600
Qy 879 TTCGGCTGATGATGCTGATCTGTGGGCAAGTTTGTCTGATACAAATATGTCACCTGT 938
Db 601 TTCGGCTGATGATGCTGATCTGTGGGCAAGTTTGTCTGATACAAATATGTCACCTGT 660
Qy 939 TGGCTGTGACAGAAAGATGATGATTTTGAAGGCGCTGGCTTCAATGCTTTGAAGA 998
Db 661 TGGCTGTGACAGAAAGATGATGATTTTGAAGGCGCTGGCTTCAATGCTTTGAAGA 720
Qy 999 AAGGCGAAGCAAGTGGGATGCTGTGCCAACATGAAGTGTGCTTTTGAACAAC 1058
Db 721 AAGGCGAAGCAAGTGGGATGCTGTGCCAACATGAAGTGTGCTTTTGAACAAC 780
Qy 1059 CCGCGCTGACAGGACCACTTGGCTCATTCACATCAACACCAAGCGCTGGTGGCGCGC 1118
Db 781 CCGCGCTGACAGGACCACTTGGCTCATTCACATCAACACCAAGCGCTGGTGGCGCGC 840
Qy 1119 TACATCTTCAAGACTCAAGTTCCTTGAATAAAGAACTCTGAGGCTCTCTGCTT 1178
Db 841 TACATCTTCAAGACTCAAGTTCCTTGAATAAAGAACTCTGAGGCTCTCTGCTT 900
Qy 1179 CTATTCCTGCGCTCTGAGCAAGGCTGCACTAGGATACCTGCTGCTTCCAGATGGA 1238
Db 901 CTATTCCTGCGCTCTGAGCAAGGCTGCACTAGGATACCTGCTGCTTCCAGATGGA 960
Qy 1239 TTCTCATTTGTTATTTGGAAGAACAGAGCTGCCAGGCTCATTCAGAGAGCCCACTTG 1298
Db 961 TTCTCATTTGTTATTTGGAAGAACAGAGCTGCCAGGCTCATTCAGAGAGCCCACTTG 1020
Qy 1299 AGCAAGCTGGCGCATTAAGTCTCTCCAGCCCTTCTACTATTTGCTGCAACAGACCATC 1358
Db 1021 AGCAAGCTGGCGCATTAAGTCTCTCCAGCCCTTCTACTATTTGCTGCAACAGACCATC 1080
Qy 1359 CACTGCTCTTCAATGGTTAATCTCAAGACTGCTTGTGCTCTTCAAGTGGGCAATATG 1418
Db 1081 CACTGCTCTTCAATGGTTAATCTCAAGACTGCTTGTGCTCTTCAAGTGGGCAATATG 1140
Qy 1419 CTTAAGGTGATTAATCAATCTATTTCTTGGCCACATCTTCTCTGAGCTACTATTC 1478
Db 1141 CTTAAGGTGATTAATCAATCTATTTCTTGGCCACATCTTCTCTGAGCTACTATTC 1200
Qy 1479 ATATGCTTTATTAATCAAGAAAGATGTCAGAAAGAGAAATTAAGAAAGATGAA 1538
Db 1201 ATATGCTTTATTAATCAAGAAAGATGTCAGAAAGAGAAATTAAGAAAGATGAA 1260
Qy 1539 TAATCAATTTCCCTGAGTGTGGGCACTGAGGAGCTGAGCAAAATACTGCTCCCTTTC 1598
Db 1261 TAATCAATTTCCCTGAGTGTGGGCACTGAGGAGCTGAGCAAAATACTGCTCCCTTTC 1320
Qy 1599 ACAGCACTCTTTGCCCGAGAGCAAGAAATGAAAGCAAGGAGTGAAGATGCAATGC 1658
Db 1321 ACAGCACTCTTTGCCCGAGAGCAAGAAATGAAAGCAAGGAGTGAAGATGCAATGC 1380
Qy 1659 TTCAAGTGTGCTCTGCTGCGCAGCAAGTCTTATTTGGGCGCAAGGGGAAATCTTTT 1718
Db 1381 TTCAAGTGTGCTCTGCTGCGCAGCAAGTCTTATTTGGGCGCAAGGGGAAATCTTTT 1440
Qy 1719 TTTGGAAGAGGCGCTTGTGTTGTCACCAAGCTGGAATGAGTGGGAGATCTCAGCTC 1778
Db 1441 TTTGGAAGAGGCGCTTGTGTTGTCACCAAGCTGGAATGAGTGGGAGATCTCAGCTC 1500
Qy 1779 ACCGCAACTCTCACTCTGAGGTTCAAGTATTTCTGCTCAGGCTCCCAAGTATGCTG 1838

Db	1501	ACCGAACTCACCTCTGGGTTCAATGATTTTTCGCTCAAGCTCCCAATGACTG	1566
Qy	1839	GGATATACGGCACTGCAACCATATGCTCCAGCTAATTTTTGTATTTTCATGTAAGAAACGGGATTT	18999
Db	1561	GGAAATACGGGACCGCACCATATGCCAGGATTAATTTTTGTATTTTCATGTAAGAAACGGGATTT	1620
Qy	1899	CACCACTGTGGCCAGAGCTGATCTCGAATCTCTGACCGCAATGATATCCACCGCTCCGCC	1958
Db	1621	CACCACTGTGGCCAGAGCTGATCTCGAATCTCTGACCGCAATGATATCCACCGCTCCGCC	1680
Qy	1959	TCCCAAAATGCTGGGATTTACAGAGCGTGAAGCCACGCTGCCCGGCCCAAGGGGAAACTTT	2018
Db	1661	TCCCAAAATGCTGGGATTTACAGAGCGTGAAGCCACGCTGCCCGGCCCAAGGGGAAACTTT	1740
Qy	2019	GTGGAGAGAGCAGAGGGGCTCAACATCTCCCTCTGATTCGCCCATGACATTTGCCTTATC	2078
Db	1741	GTGGAGAGAGCAGAGGGGCTCAACATCTCCCTCTGATTCGCCCATGACATTTGCCTTATC	1800
Qy	2079	TCTCCCATCTAGCCAGGAATCTATGTGTTTTCTTCTGCG	2120
Db	1801	TCTCCCATCTAGCCAGGAATCTATGTGTTTTCTTCTGCG	1842

```

RESULT 4
US-09-764-860-1135/C
; Sequence 1135, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1135
; LENGTH: 7461
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-860-1135

Query Match          31.2%; Score 708.6; DB 9; Length 7461;
Best Local Similarity 99.4%; Pred. No. 2.8e-209;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1546 TTTCCTGTGTGGCTGTGCGGGAGCTGTGTGCAGAAATCTACTGCTCTCCCTTTTCACAGCAC 1605
DB 721 TCTTGCAAGTGGCTGTGCGGGAGCTGTGTGCAGAAATCTACTGCTCTCCCTTTTCACAGCAC 662

OY 1606 TCTTTGGCCCAAGACAGAGATGGAAGAAGCCAGGAGGTGGAGATGATGATCTTCACGC 1665
DB 661 TCTTTGGCCCAAGACAGAGATGGAAGAAGCCAGGAGGTGGAGATGATGATCTTCACGC 602

OY 1666 TGTGCTCTGTGTCAGCCAGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTTTTTGGAG 1725
DB 601 TGTGCTCTGTGTCAGCCAGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTTTTTGGAG 542

OY 1726 AAGGGGTCTTGTCTTTGTATCCACCGCTGGATGATGATGTCGCGGATCTCACTCACCGCAA 1785
DB 541 AAGGGGTCTTGTCTTTGTATCCACCGCTGGATGATGATGTCGCGGATCTCACTCACCGCAA 482

OY 1786 CCTCACCTCTCGGGGTTCAAGTATTTTTCCTGACCTCCAGCTCCCAAGTAGCTGGGAAATAC 1845
DB 481 CCTCACCTCTCGGGGTTCAAGTATTTTTCCTGACCTCCAGCTCCCAAGTAGCTGGGAAATAC 422

OY 1846 AAGCACGCCACCATGCCAGCTAATTTTTGTATTTTTCAGTAGAAACGGGATTTCCACACG 1905
DB 421 AAGCACGCCACCATGCCAGCTAATTTTTGTATTTTTCAGTAGAAACGGGATTTCCACACG 362

OY 1906 TTGGCAGAGCTGTGTTGCAATCTCTGACACCGCAAGATTCACACCGCCTCTCGCTCTCCAAA 1965
DB 361 TTGGCAGAGCTGTGTTGCAATCTCTGACACCGCAAGATTCACACCGCCTCTCGCTCTCCAAA 302

```

QY	1966	GTGCTGGGATTATACAGGGGTAGCCACCGTCGCCGGCCCAAAAGGGGAAACTCTTGTTGGAG	2025
Db	301	GTGCTGGGATTATACAGGCGTGAAGCCACCGTGCCGGCCCAAAAGGGGAAACTCTTGTTGGAG	242
QY	2026	GAGCAGAGAGGGCTCACATCTCCCTCTGATTCCCCCATGACCATTTGCTTATCTCTCCC	2085
Db	241	GAGCAGAGAGGGCTCACATCTCCCTCTGATTCCCCCATGACCATTTGCTTATCTCTCCC	182
QY	2086	ATTACAGCAGGAATCTATATGTGTTTCTTCTGCAATTATCTATGATGTGATGTGCC	2145
Db	181	ATTACAGCAGGAATCTATATGTGTTTCTTCTGCAATTATCTATGATGTGATGTGCC	122
QY	2146	GCTACCAACCAACCCCCCATGAGGGGGGTGAGAGGGGTGCAAAGGCCCTGCTGCTCACT	2205
Db	121	GCTACCAACCAACCCCCCATGAGGGGGGTGAGAGGGGTGCAAAGGCCCTGCTGCTCACT	62
QY	2206	TTTTCTACCTTGGAACTGTATTATATATAATACATCTCTGTTGTCTCACTTTTTC	2260
Db	61	TTTTCTACCTTGGAACTGTATTATATATAATACATCTCTGTTGTCTCACTTTTTC	7

```

RESULT 5
US-09-764-904-91/c
; Sequence 91, Application US/09764904
; Patent No. US2002017345A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764,904
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 91
; LENGTH: 7461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-904-91

Query Match          31.2%; Score 708.6; DB 10; Length 7461;
Best Local Similarity 99.4%; Pred. No. 2.8e-203;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY      1546  TTTCCTTGCTGGCTGCTGTCGGGAGCTGCGAGAAACTACGTCCTCCCTTTTCACAGAC 1605
DB      721  TCTTGCAAGTGGCTGTGCGGAGCTGTGCGGAGCTGTGCGAGAACTACTGCTCTTTCACAGAC 662

QY      1606  TCCTTTGGCCACAGACAGAGATGAGAAAGCCAGGAGGTGAGAGATCGATGCTTTCAGC 1665
DB      661  TCCTTTGGCCACAGACAGAGATGAGAAAGCCAGGAGGTGAGAGATCGATGCTTTCAGC 602

QY      1666  TGTGCTCTGCTGCGCAGCAGCAAGTCTTCACTTTGGGGCCAAAGGGGAACTTTTTTTGGAG 1722
DB      601  TGTGCTCTGCTGCGCAGCAGCAAGTCTTCACTTTGGGGCCAAAGGGGAACTTTTTTTGGAG 542

QY      1726  AAGGGCTTGTCTTTGTTCACCCACGCTGAGAAATGACATGAGCGGGATCTCAGCTCACCGCA 1789
DB      541  AAGGGCTTGTCTTTGTTCACCCACGCTGAGAAATGACATGAGCGGGATCTCAGCTCACCGCA 482

QY      1786  CCTCACCTCTCTGGGTTCAAGTGAATTTTCTGCTCAGCCTTCCCAAGTAGCTGGGAAATAC 1845
DB      481  CCTCACCTCTCTGGGTTCAAGTGAATTTTCTGCTCAGCCTTCCCAAGTAGCTGGGAAATAC 422

QY      1846  AGGCAAGCCACCAATGCCCCAAGCTAATTTTGTATTTTCAATGAGAAACGGGATTTCAACACG 1905
DB      421  AGGCAAGCCACCAATGCCCCAAGCTAATTTTGTATTTTCAATGAGAAACGGGATTTCAACACG 362

QY      1906  TTGGCAGAGCTGTGTGTGCAATCTCCAGACCGCAAGTATCCACCGGCTCCGGCTCCCAAA 1965
DB      361  TTGGCAGAGCTGTGTGTGCAATCTCCAGACCGCAAGTATCCACCGGCTCCGGCTCCCAAA 302

```


PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/236,369
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/224,519
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/220,964
 PRIOR FILING DATE: 2000-07-26
 PRIOR APPLICATION NUMBER: 60/241,809
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/249,299
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/236,327
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/241,785
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/244,617
 PRIOR FILING DATE: 2000-11-01
 PRIOR APPLICATION NUMBER: 60/225,268
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/236,368
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/251,856
 PRIOR FILING DATE: 2000-12-08
 PRIOR APPLICATION NUMBER: 60/251,868
 PRIOR FILING DATE: 2000-12-08
 PRIOR APPLICATION NUMBER: 60/229,344
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/234,997
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: 60/229,343
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/229,345
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/229,287
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/229,513
 PRIOR FILING DATE: 2000-09-05
 PRIOR APPLICATION NUMBER: 60/231,413
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/229,509
 PRIOR FILING DATE: 2000-09-05
 PRIOR APPLICATION NUMBER: 60/236,367
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/237,039
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,038
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/236,370
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/236,802
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,037
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,040
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/240,960
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/239,935
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/239,937
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/241,787
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/246,474
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/246,532
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/249,216
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,210
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/226,681
 PRIOR FILING DATE: 2000-08-22

PRIOR APPLICATION NUMBER: 60/225,759
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/225,213
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/227,182
 PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 60/225,214
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/235,836
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: 60/230,438
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/215,135
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: 60/225,266
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/249,218
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,208
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,213
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,212
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,207
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,245
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,244
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,217
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,211
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,215
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,264
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,214
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,297
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/232,400
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/231,242
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,081
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,080
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,414
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,244
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/233,064
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/233,063
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,397
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,399
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,401
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/241,808
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,826
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,786
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,221
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/246,475

```
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match      31.2%; Score 708.6; DB 14; Length 7461;
Best Local Similarity 99.4%; Pred. No. 2,8e-209;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1546 TTTCCCTGGTGGCTGTGGGGGACTGTGTGAGAACTACTGCTCTCTTTTACAGCAC 1605
DB 721 TCTTGGAGGTGGCTGTGGGGGCTGTGTGAGAACTACTGCTCTCTTTTACAGCAC 662
QY 1606 TCTTTTGGCCCAAGAGAGAAATGAAAAGCCAGGAGGAGTGAATGCTTCACAG 1665
DB 661 TCTTTTGGCCCAAGAGAGAAATGAAAAGCCAGGAGGAGTGAATGCTTCACAG 602
QY 1666 TGTGCTCTGTGCTGCAAGCTTCTTATTTGGGGCCAAAGGGGAACTTTTTTGGAG 1725
DB 601 TGTGCTCTGTGCTGCAAGCTTCTTATTTGGGGCCAAAGGGGAACTTTTTTGGAG 542
QY 1726 AAGGCTCTGTGCTTGTGTACCCACGCTGGAATGAGTGGGGGATCTCAGTCAAGCA 1785
DB 541 AAGGCTCTGTGCTTGTGTACCCACGCTGGAATGAGTGGGGGATCTCAGTCAAGCA 482
QY 1786 CCTCCACCTCTCTGGGTTCAAGTATTTCTGCTCAGCTCCCAAGTACGAGGAAATAC 1845
DB 481 CCTCCACCTCTCTGGGTTCAAGTATTTCTGCTCAGCTCCCAAGTACGAGGAAATAC 422
QY 1846 AAGGACGCGACATGCGCCAGCTAATTTTGTATTTTCAAGTAAAGGGAATTCACACAG 1905
DB 421 AAGGACGCGACATGCGCCAGCTAATTTTGTATTTTCAAGTAAAGGGAATTCACACAG 362
QY 1906 TTGGCCAGGCTGTGTGCAATCTCTGACCGCAAGTATCACCGGCTCTCGCTCCCAAA 1965
DB 361 TTGGCCAGGCTGTGTGCAATCTCTGACCGCAAGTATCACCGGCTCTCGCTCCCAAA 302
QY 1966 GTGCTGGGATTACAGCGGTGAGCCAGCGGCCGCAAGGGGAACTCTGTGGAG 2025
DB 301 GTGCTGGGATTACAGCGGTGAGCCAGCGGCCGCAAGGGGAACTCTGTGGAG 242
QY 2026 GAGCAGAGGGGCTCACATCTCCCTCTGATTTCCCAATGACATTCCTTATCTCTCCC 2085
DB 241 GAGCAGAGGGGCTCACATCTCCCTCTGATTTCCCAATGACATTCCTTATCTCTCCC 182
QY 2086 ATCTACCGAGGATCTATTTGTCTTCTTCCCAATTTACTATGATTTGATGTGCC 2145
DB 181 ATCTACCGAGGATCTATTTGTCTTCTTCCCAATTTACTATGATTTGATGTGCC 122
QY 2146 GCTACACACACCCCCCATGSGGGGTGAGAGGGGTGCAAGGCCCTGCTCTCCACT 2205
DB 121 GCTACACACACCCCCCATGSGGGGTGAGAGGGGTGCAAGGCCCTGCTCTCCACT 62
QY 2206 TTTTCTACCTTGGAAGTATTAATAATCACTTCTGTTTGTCAAGTTTCA 2260
DB 61 TTTTCTACCTTGGAAGTATTAATAATCACTTCTGTTTGTCAAGTTTCA 7

RESULT 8
US-09-908-711-158/c
; Sequence 158, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA128
; CURRENT APPLICATION NUMBER: US/09/908,711
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US01/01360
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,867
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01344
; PRIOR FILING DATE: 2001-01-17
```

```
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01345
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,905
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01354
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01339
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01340
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,874
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01334
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,898
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01320
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,853
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01349
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,902
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01239
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01348
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,882
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01347
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,896
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01307
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01341
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,856
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01336
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,868
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01312
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 158
; LENGTH: 32174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```


NAME/KEY: SITE	LOCATION: (29356)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29357)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29358)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29359)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29360)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29361)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29362)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29363)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29364)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29365)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29366)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29367)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29368)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29369)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29370)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29371)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29372)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29373)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29374)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29375)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29376)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29377)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29378)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29379)	n equals a,t,g, or c
OTHER INFORMATION: NAME/KEY: SITE	LOCATION: (29380)	n equals a,t,g, or c

1	LOCATION: (29380)	n	equals	a,t,g,	or c
2	OTHER INFORMATION:	n	equals	a,t,g,	or c
3	NAME/KEY: SITE	n	equals	a,t,g,	or c
4	LOCATION: (29381)	n	equals	a,t,g,	or c
5	OTHER INFORMATION:	n	equals	a,t,g,	or c
6	NAME/KEY: SITE	n	equals	a,t,g,	or c
7	LOCATION: (29382)	n	equals	a,t,g,	or c
8	OTHER INFORMATION:	n	equals	a,t,g,	or c
9	NAME/KEY: SITE	n	equals	a,t,g,	or c
10	LOCATION: (29383)	n	equals	a,t,g,	or c
11	OTHER INFORMATION:	n	equals	a,t,g,	or c
12	NAME/KEY: SITE	n	equals	a,t,g,	or c
13	LOCATION: (29384)	n	equals	a,t,g,	or c
14	OTHER INFORMATION:	n	equals	a,t,g,	or c
15	NAME/KEY: SITE	n	equals	a,t,g,	or c
16	LOCATION: (29385)	n	equals	a,t,g,	or c
17	OTHER INFORMATION:	n	equals	a,t,g,	or c
18	NAME/KEY: SITE	n	equals	a,t,g,	or c
19	LOCATION: (29386)	n	equals	a,t,g,	or c
20	OTHER INFORMATION:	n	equals	a,t,g,	or c
21	NAME/KEY: SITE	n	equals	a,t,g,	or c
22	LOCATION: (29387)	n	equals	a,t,g,	or c
23	OTHER INFORMATION:	n	equals	a,t,g,	or c
24	NAME/KEY: SITE	n	equals	a,t,g,	or c
25	LOCATION: (29388)	n	equals	a,t,g,	or c
26	OTHER INFORMATION:	n	equals	a,t,g,	or c
27	NAME/KEY: SITE	n	equals	a,t,g,	or c
28	LOCATION: (29389)	n	equals	a,t,g,	or c
29	OTHER INFORMATION:	n	equals	a,t,g,	or c
30	NAME/KEY: SITE	n	equals	a,t,g,	or c
31	LOCATION: (29390)	n	equals	a,t,g,	or c
32	OTHER INFORMATION:	n	equals	a,t,g,	or c
33	NAME/KEY: SITE	n	equals	a,t,g,	or c
34	LOCATION: (29391)	n	equals	a,t,g,	or c
35	OTHER INFORMATION:	n	equals	a,t,g,	or c
36	NAME/KEY: SITE	n	equals	a,t,g,	or c
37	LOCATION: (29392)	n	equals	a,t,g,	or c
38	OTHER INFORMATION:	n	equals	a,t,g,	or c
39	NAME/KEY: SITE	n	equals	a,t,g,	or c
40	LOCATION: (29393)	n	equals	a,t,g,	or c

Query Match	31.2%	Score 708.6;	DB 9;	Length 32174;
Best Local Similarity	99.4%;	Pred. No. 7.2e-209;		
Matches 711; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1546	TTTTCCCTGGTGGCCCTGTGCGGGGACCTGGGTCAGAAACTATCGTCTCCCTTTTTCACAGAC	1605
Db	6050	TCCTTCAGAGTGGCTGTGCGGGACTGTGCAGAAACTATCGTCTCCCTTTTTCACAGAC	5991
QY	1606	TCCTTTGGCCCCAGAGCAGAGAAATGMAAAGCCAGGAGAGTGGAAAGATGATGCTTCCAGC	1665
Db	5990	TCCTTTGGCCCCAGAGCAGAGAAATGMAAAGCCAGGAGAGTGGAAAGATGATGCTTCCAGC	5933
QY	1666	TGTGCTCTGTGCTGCGCAGCCAAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTTTTGGAG	1725
Db	5930	TGTGCTCTGTGCTGCGCAGCCAAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTTTTGGAG	5873
QY	1726	AAGGCGCTTGCTTTGTTCACCCACGCGTGGAAATGCAATGGCGGGAACTCAGCTCACCGGA	1785
Db	5870	AAGGCGCTTGCTTTGTTCACCCACGCGTGGAAATGCAATGGCGGGAACTCAGCTCACCGGA	5811
QY	1786	CCTCCACCTCCTCGGGTTCAAGTGAATTTTTCCTGACTCAGCTCCGCCAAGTAGCTGGGAATAC	1845
Db	5810	CCTCCACCTCCTCGGGTTCAAGTGAATTTTTCCTGACTCAGCTCCGCCAAGTAGCTGGGAATAC	5755
QY	1846	AGGCAACGCGCACCATGCCCCAGCTAAATTTTGTATTTTCAGTAGAAACGGGAATTTACACAG	1905
Db	5750	AGGCAACGCGCACCATGCCCCAGCTAAATTTTGTATTTTCAGTAGAAACGGGAATTTACACAG	5691
QY	1906	TTGGCCAGAGCGGTGCTGGAACTCCTGACCGCAAGTAGTCCACCCGCTCCGCGCTCCCAAA	1965
Db	5690	TTGGCCAGAGCGGTGCTGGAACTCCTGACCGCAAGTAGTCCACCCGCTCCGCGCTCCCAAA	5633

[illegible]

1	OTHER INFORMATION:	n	equals	a,t,g,	or	c
2	NAME/KEY: SITE					
3	LOCATION: (29392)					
4	OTHER INFORMATION:	n	equals	a,t,g,	or	c
5	NAME/KEY: SITE					
6	LOCATION: (29393)					
7	OTHER INFORMATION:	n	equals	a,t,g,	or	c
8	NAME/KEY: SITE					
9	LOCATION: (29394)					
10	OTHER INFORMATION:	n	equals	a,t,g,	or	c
11	NAME/KEY: SITE					
12	LOCATION: (29395)					
13	OTHER INFORMATION:	n	equals	a,t,g,	or	c
14	NAME/KEY: SITE					
15	LOCATION: (29396)					
16	OTHER INFORMATION:	n	equals	a,t,g,	or	c
17	NAME/KEY: SITE					
18	LOCATION: (29397)					
19	OTHER INFORMATION:	n	equals	a,t,g,	or	c
20	NAME/KEY: SITE					
21	LOCATION: (29398)					
22	OTHER INFORMATION:	n	equals	a,t,g,	or	c
23	NAME/KEY: SITE					
24	LOCATION: (29399)					
25	OTHER INFORMATION:	n	equals	a,t,g,	or	c
26	NAME/KEY: SITE					
27	LOCATION: (29400)					
28	OTHER INFORMATION:	n	equals	a,t,g,	or	c
29	NAME/KEY: SITE					
30	LOCATION: (29401)					
31	OTHER INFORMATION:	n	equals	a,t,g,	or	c
32	NAME/KEY: SITE					
33	LOCATION: (29402)					
34	OTHER INFORMATION:	n	equals	a,t,g,	or	c
35	NAME/KEY: SITE					
36	LOCATION: (29403)					
37	OTHER INFORMATION:	n	equals	a,t,g,	or	c
38	NAME/KEY: SITE					
39	LOCATION: (29404)					
40	OTHER INFORMATION:	n	equals	a,t,g,	or	c
41	NAME/KEY: SITE					
42	LOCATION: (29405)					
43	OTHER INFORMATION:	n	equals	a,t,g,	or	c
44	NAME/KEY: SITE					
45	LOCATION: (29406)					
46	OTHER INFORMATION:	n	equals	a,t,g,	or	c
47	NAME/KEY: SITE					
48	LOCATION: (29407)					
49	OTHER INFORMATION:	n	equals	a,t,g,	or	c
50	NAME/KEY: SITE					
51	LOCATION: (29408)					
52	OTHER INFORMATION:	n	equals	a,t,g,	or	c
53	NAME/KEY: SITE					
54	LOCATION: (29409)					
55	OTHER INFORMATION:	n	equals	a,t,g,	or	c
56	NAME/KEY: SITE					
57	LOCATION: (29410)					
58	OTHER INFORMATION:	n	equals	a,t,g,	or	c
59	NAME/KEY: SITE					
60	LOCATION: (29411)					
61	OTHER INFORMATION:	n	equals	a,t,g,	or	c
62	NAME/KEY: SITE					
63	LOCATION: (29412)					
64	OTHER INFORMATION:	n	equals	a,t,g,	or	c
65	NAME/KEY: SITE					
66	LOCATION: (29413)					
67	OTHER INFORMATION:	n	equals	a,t,g,	or	c
68	NAME/KEY: SITE					
69	LOCATION: (29414)					
70	OTHER INFORMATION:	n	equals	a,t,g,	or	c
71	NAME/KEY: SITE					
72	LOCATION: (29415)					
73	OTHER INFORMATION:	n	equals	a,t,g,	or	c

```

; NAME/KEY: SITE
; LOCATION: (29416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

```

Query Match	31.2%;	Score 708.6;	DB 9;	Length 32174;
Best Local Similarity	99.4%;	Pred. No. 7.2e-209;		
Matches 711; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

Qy	1546	TTTTCCCTGGTGGCCCTGTGCGGGAATGGGTGCAGAAACTATGTCCTCCCTTTTCACAGCAC	1600
Db	6050	TCTTGACAGTGGCCCTGTGCGGGAATGGGTGCAGAAACTATGTCCTCCCTTTTCACAGCAC	5999
Qy	1606	TCTCTTTGCCCCAGACGACAGAAATGGAAAAAGCCAGGAGGTGGAAATGATGATGCTTCCACAG	1666
Db	5990	TCTCTTTGCCCCAGACGACAGAAATGGAAAAAGCCAGGAGGTGGAAATGATGATGCTTCCACAG	5933
Qy	1666	TGTGCTCTGTGTCGCACACCAAGTCTTGCAATTTGGGGGCAAAAGGGGAAACTTTTCTTTTGGAG	1726
Db	5930	TGTGCTCTGTGTCGCACACCAAGTCTTGCAATTTGGGGGCAAAAGGGGAAACTTTTCTTTTGGAG	5877
Qy	1726	AAGCGCTCTTGCTTTGTTCACCCACGCTGGAAATGCAATGGCGGGAATCTCAGCTCACCGCA	1788
Db	5870	AAGCGCTCTTGCTTTGTTCACCCACGCTGGAAATGCAATGGCGGGAATCTCAGCTCACCGCA	5811
Qy	1786	CCTCCACCTCCTGGGGTTCAGATGATTTTCCGCGCCACAGCCCTCCCAAGTAAGCGGGGAATAC	1846
Db	5810	CCTCCACCTCCTGGGGTTCAGATGATTTTCCGCGCCACAGCCCTCCCAAGTAAGCGGGGAATAC	5755
Qy	1846	AGGCAAGCCCAACCATGCCAGCTAATTTTGTATTTTCAGTAGAAACGGGATTTTCCACAG	1906
Db	5750	AGGCAAGCCCAACCATGCCAGCTAATTTTGTATTTTCAGTAGAAACGGGATTTTCCACAG	5699
Qy	1906	TTGGCCAGGCTGGTCTTGAACTCCTGACCGGAATGCACCCGCTCCGCTCCCAAA	1966
Db	5690	TTGGCCAGGCTGGTCTTGAACTCCTGACCGGAATGCACCCGCTCCGCTCCCAAA	5633
Qy	1966	GTGCTGGGATTTACAGGCGGTGAGCCACGCGTCCCGGCCCAAAAGGGGAAACTCTGTGGAG	2026
Db	5630	GTGCTGGGATTTACAGGCGGTGAGCCACGCGTCCCGGCCCAAAAGGGGAAACTCTGTGGAG	5577
Qy	2026	GAGCAGAGGGGCTCACTCTCCCTCTGATTTCCCCATGCACATTTGCCTTATCTCTCCC	2086
Db	5570	GAGCAGAGGGGCTCACTCTCCCTCTGATTTCCCCATGCACATTTGCCTTATCTCTCCC	5511
Qy	2086	ATCAGACAGGAATCTATTTGTGTTTTCTTGCCAAATTACATGATGTGTGATGTGTC	2146
Db	5510	ATCAGACAGGAATCTATTTGTGTTTTCTTGCCAAATTACATGATGTGTGATGTGTC	5455
Qy	2146	GCTACCAACACCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTGTCACT	2206
Db	5450	GCTACCAACACCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTGTCACT	5399
Qy	2206	TTTTCTACCTTTGGAACTGTATTAATAAACAATCTGTTTGTTCAGATTTTTCA	2260
Db	5390	TTTTCTACCTTTGGAACTGTATTAATAAACAATCTGTTTGTTCAGATTTTTCA	5336

```

RESULT 10
US-09-764-877-2645/c
Sequence 2645, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC0005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2645
LENGTH: 32174

```


[illegible]

```

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

Query Match      31.2%; Score 708.6; DB 10; Length 32174;
Best Local Similarity 99.4%; Pred. No. 7.2e-209;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1546 TTTCCCTGGTGGCTGTGGGGGAGCTGTGCAAACTACTGCTCTCTTTTTCACAGCAC 1605
DB 6050 TCTTGAGGTGGCTGTGGGGGAGCTGTGCAAACTACTGCTCTCTTTTTCACAGCAC 5991
QY 1606 TCTTTGGCCCCAGAGAGAAATGAGAAAGCCAGGAGGAGAGATGCAATGCTTCACAC 1665
DB 5990 TCTTTGGCCCCAGAGAGAAATGAGAAAGCCAGGAGGAGAGATGCAATGCTTCACAC 5931
QY 1666 TGTGCTCTGTGTCAGCCAGCAAGCTTTCATTTGGGGCCAAAGGGGAACTTTTGTGAG 1725
DB 5930 TGTGCTCTGTGTCAGCCAGCAAGCTTTCATTTGGGGCCAAAGGGGAACTTTTGTGAG 5871
QY 1726 AAGGCTCTTGTCTTTGTTCACCCAGCTGGAATGAGTGGGGATCTCAGCTCACGGCAA 1785
DB 5870 AAGGCTCTTGTCTTTGTTCACCCAGCTGGAATGAGTGGGGATCTCAGCTCACGGCAA 5811
QY 1786 CCTCCACCTCCTGGGGTTCAAGATTTTCTGCTGAGCTCCCAAGTACCTGGGAATAC 1845
DB 5810 CCTCCACCTCCTGGGGTTCAAGATTTTCTGCTGAGCTCCCAAGTACCTGGGAATAC 5751
QY 1846 AAGCAAGCCACCAATGCCAGCACTAATTTTGTATTTTCAGTAGAAAGGATTTTCACAGC 1905
DB 5750 AAGCAAGCCACCAATGCCAGCACTAATTTTGTATTTTCAGTAGAAAGGATTTTCACAGC 5691
QY 1906 TTGGCCAGGCTGTCTGGAATCTCTGACCGCAAGTATCCACCCGCTCCGCCCAA 1965
DB 5690 TTGGCCAGGCTGTCTGGAATCTCTGACCGCAAGTATCCACCCGCTCCGCCCAA 5631
QY 1966 GTGCTGGGATTACAGGCGTAGAGCAACCGGCCCGGCCAAAGGGGAACTTTGTGGAG 2025
DB 5630 GTGCTGGGATTACAGGCGTAGAGCAACCGGCCCGGCCAAAGGGGAACTTTGTGGAG 5571
QY 2026 GAGCAGAGGGGCTCATATCTCCCTCTGATTTCCCCATGACATTCCTTATCTCCCC 2085
DB 5570 GAGCAGAGGGGCTCATATCTCCCTCTGATTTCCCCATGACATTCCTTATCTCCCC 5511
QY 2086 ATCTAGCCAGGAATCTATTGTGTTTCTTCTGCAATTTACTATGATGTGATGTC 2145
DB 5510 ATCTAGCCAGGAATCTATTGTGTTTCTTCTGCAATTTACTATGATGTGATGTC 5451
QY 2146 GGTACCAACACCCCCCCCATTTGGGGGGTGGAGAGGGGTGCAAGCCCTGCTGCACT 2205
DB 5450 GGTACCAACACCCCCCCCATTTGGGGGGTGGAGAGGGGTGCAAGCCCTGCTGCACT 5391
QY 2206 TTTTCTACCTTGAATCTATTAGATAAATCACTCTGTTGTTGCAATTTTCA 2260
DB 5390 TTTTCTACCTTGAATCTATTAGATAAATCACTCTGTTGTTGCAATTTTCA 5336

RESULT 12
US-09-764-904-90/c.
; Sequence 90, Application US/09764904
; Patent No. US20020173454A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764,904
; PRIORITY FILING DATE: 2001-01-17
; PRIORITY APPLICATION DATA REMOVED - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 90
```

```

; LENGTH: 32174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (29356)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29359)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29360)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29361)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29362)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29363)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29364)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29365)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29366)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29367)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29368)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29369)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29370)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29371)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29372)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29373)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29374)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29375)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29376)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29378)
; OTHER INFORMATION: n equals a,t,g, or c
```

NAME/KEY: SITE	
LOCATION: (29379)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29380)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29381)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29382)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29383)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29384)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29385)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29386)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29387)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29388)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29389)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29390)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29391)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29392)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29393)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29394)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29395)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29396)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29397)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29398)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29399)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29400)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29401)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	
LOCATION: (29402)	n equals a,t,g, or c
OTHER INFORMATION:	
NAME/KEY: SITE	

1	LOCATION: (29403)	n	equals	a,t,g,	or c
2	OTHER INFORMATION:	n	equals	a,t,g,	or c
3	NAME/KEY: SITE	n	equals	a,t,g,	or c
4	LOCATION: (29406)	n	equals	a,t,g,	or c
5	OTHER INFORMATION:	n	equals	a,t,g,	or c
6	NAME/KEY: SITE	n	equals	a,t,g,	or c
7	LOCATION: (29405)	n	equals	a,t,g,	or c
8	OTHER INFORMATION:	n	equals	a,t,g,	or c
9	NAME/KEY: SITE	n	equals	a,t,g,	or c
10	LOCATION: (29406)	n	equals	a,t,g,	or c
11	OTHER INFORMATION:	n	equals	a,t,g,	or c
12	NAME/KEY: SITE	n	equals	a,t,g,	or c
13	LOCATION: (29407)	n	equals	a,t,g,	or c
14	OTHER INFORMATION:	n	equals	a,t,g,	or c
15	NAME/KEY: SITE	n	equals	a,t,g,	or c
16	LOCATION: (29408)	n	equals	a,t,g,	or c
17	OTHER INFORMATION:	n	equals	a,t,g,	or c
18	NAME/KEY: SITE	n	equals	a,t,g,	or c
19	LOCATION: (29409)	n	equals	a,t,g,	or c
20	OTHER INFORMATION:	n	equals	a,t,g,	or c
21	NAME/KEY: SITE	n	equals	a,t,g,	or c
22	LOCATION: (29410)	n	equals	a,t,g,	or c
23	OTHER INFORMATION:	n	equals	a,t,g,	or c
24	NAME/KEY: SITE	n	equals	a,t,g,	or c
25	LOCATION: (29411)	n	equals	a,t,g,	or c
26	OTHER INFORMATION:	n	equals	a,t,g,	or c
27	NAME/KEY: SITE	n	equals	a,t,g,	or c
28	LOCATION: (29412)	n	equals	a,t,g,	or c
29	OTHER INFORMATION:	n	equals	a,t,g,	or c
30	NAME/KEY: SITE	n	equals	a,t,g,	or c
31	LOCATION: (29413)	n	equals	a,t,g,	or c
32	OTHER INFORMATION:	n	equals	a,t,g,	or c
33	NAME/KEY: SITE	n	equals	a,t,g,	or c
34	LOCATION: (29414)	n	equals	a,t,g,	or c
35	OTHER INFORMATION:	n	equals	a,t,g,	or c
36	NAME/KEY: SITE	n	equals	a,t,g,	or c
37	LOCATION: (29415)	n	equals	a,t,g,	or c
38	OTHER INFORMATION:	n	equals	a,t,g,	or c
39	NAME/KEY: SITE	n	equals	a,t,g,	or c
40	LOCATION: (29416)	n	equals	a,t,g,	or c
41	OTHER INFORMATION:	n	equals	a,t,g,	or c
42	NAME/KEY: SITE	n	equals	a,t,g,	or c

Query Match	31.2%;	Score 708.6;	DB 10;	Length 32174;
Best Local Similarity	99.4%;	Pred. No. 7.2e-209;		
Matches 711;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

Oy	1546	TTTCCCTGAGTGGCGCTGTGACGGGAACTGGGAGCAAACTACTCGTCTCCCTTTTCAAGAC	1605
Db	6050	TCTTGACAGTGGCCCTGTGCGGGAACTGGGTGACGAACACTACTCGTCTCCCTTTTCAAGACAC	5991
Oy	1606	TCCTTTGCCCCAGACAGAGATGAGAAAGCCACGGAGGTGGAGAGATGATGCTTCCAGC	1665
Db	5990	TCCCTTTGCCCCAGACAGAGATGAGAAAGCCACGGAGGTGGAGAGATGATGATGCTTCCAGC	5931
Oy	1666	TGTCGCTCTGTGTGCAGCAGCAAGTCTTATATTGTTGGGGCCAAAGGGGAAACTTTTTTTGGAG	1725
Db	5930	TGTCGCTCTGTGTGCAGCAGCAAGTCTTATATTGTTGGGGCCAAAGGGGAAACTTTTTTTGGAG	5871
Oy	1726	AAGGCGTCTTGCTTTGTTCACCCACAGCTGGAAATGCGAGTGGCGGGATCTCAGCTCACCGCA	1785
Db	5870	AAGGCGTCTTGCTTTGTTCACCCACAGCTGGAAATGCGAGTGGCGGGATCTCAGCTCACCGCA	5811
Oy	1786	CCTCCACTCTCTGGGTTCAAGTATTTTCTGTGCTCAGCTCCCAAGTACTGGGAATAC	1845
Db	5810	CCTCCACTCTCTGGGTTCAAGTATTTTCTGTGCTCAGCTCCCAAGTACTGGGAATAC	5751
Oy	1846	AGGACGCCACCAATGCGCAGCTAATTTTGTATTTTCAGTAGAAACGGGATTTGACACAG	1905
Db	5750	AGGACGCCACCAATGCGCAGCTAATTTTGTATTTTCAGTAGAAACGGGATTTGACACAG	5691
Oy	1906	TTGGCCAGGCTGTGTGAACTCTTGACCGCAAGTATCCAACCGGCTCCGGCTTCCCAA	1965
Db		TTGGCCAGGCTGTGTGAACTCTTGACCGCAAGTATCCAACCGGCTCCGGCTTCCCAA	

Db 5690 TTGGCCAGGCTGCTCTCGAATCTCTGACCGCAAGTATCCACCCGCTCCGCTCCCAAA 5631
Qy 1966 GTGCTGGATTACAGGCGTAGCCACCGTGGCCCGGCAAGGGGAACTCTTGGGAG 2025
Db 5630 GTGCTGGATTACAGGCGTAGCCACCGTGGCCCGGCAAGGGGAACTCTTGGGAG 5571
Qy 2026 GAGCAGAGGGGCTCACATCTCCCTCTGATTTCCCATGACATTCCTTATCTCTCC 2085
Db 5570 GAGCAGAGGGGCTCACATCTCCCTCTGATTTCCCATGACATTCCTTATCTCTCC 5511
Qy 2086 ATCTAGCCAGATCTATTTGTTTCTTCTGCAATTTACTATATTTGATGTGCC 2145
Db 5510 ATCTAGCCAGATCTATTTGTTTCTTCTGCAATTTACTATATTTGATGTGCC 5451
Qy 2146 GCTACCAACCCCGCCCATGGGGGGGAGAGAGGGGCGAAGGCGCTGCTGCTCACT 2205
Db 5450 GCTACCAACCCCGCCCATGGGGGGGAGAGAGGGGCGTGGAGGCGCTGCTGCTCACT 5391
Qy 2206 TTTTCTACCTTGGAACTGTATTAAGATMAATCACTTCTGTTGTCAGTTTTTCA 2260
Db 5390 TTTTCTACCTTGGAACTGTATTAAGATMAATCACTTCTGTTGTTCAATTTTCA 5336

RESULT 13

US-09-764-891-6480/c
Sequence 6480, Application US/09764891
Publication No. US20030077808A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6480
LENGTH: 32174
TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE
LOCATION: (29356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29357)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29358)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29360)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29361)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29362)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29363)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29364)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29365)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29366)

OTHER INFORMATION: n equals a,t,g, or c

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29367)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29368)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29369)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29370)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29371)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29372)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29373)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29374)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29375)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29376)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29377)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29378)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29379)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29380)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29381)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29382)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29383)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29384)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29385)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29386)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29387)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29388)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29389)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29390)
OTHER INFORMATION: n equals a,t,g, or c

```

; LOCATION: (29415)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

Query Match          31.2%; Score 708.6; DB 11; Length 32174;
Best Local Similarity 99.4%; Pred. No. 7,2e-209;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1546 TTTCCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 1605
Db 6050 TCTTTCAGAGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5991

OY 1606 TCCCTTTCAGAGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 1665
Db 5990 TCCCTTTCAGAGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5991

OY 1666 TGTGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 1725
Db 5930 TGTGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5871

OY 1726 AAGCGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 1785
Db 5870 AAGCGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5811

OY 1786 CCTCAGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 1845
Db 5810 CCTCAGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5751

OY 1846 AGGAGCGTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 1905
Db 5750 AGGAGCGTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5691

OY 1906 TTGGCAGAGGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 1965
Db 5690 TTGGCAGAGGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5631

OY 1966 GTGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 2025
Db 5630 GTGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5571

OY 2026 GAGAGAGGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 2085
Db 5570 GAGAGAGGCTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5511

OY 2086 ATCTAGCAGAGTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 2145
Db 5510 ATCTAGCAGAGTGTGTGCGCTGTGTGCGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5451

OY 2146 GCTACACACACCCCCCAATGAGGAGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 2205
Db 5450 GCTACACACACCCCCCAATGAGGAGGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5391

OY 2206 TTTCTACCTGTGTGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 2260
Db 5390 TTTCTACCTGTGTGAGTGTGTGCAAGAACTACTGTCTCTCCCTTTTACAGCAC 5336

RESULT 14
US-09-764-891-10135
; Sequence 10135, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0

```

SEQ ID NO	10135
LENGTH:	32174
TYPE:	DNA
ORGANISM:	Homo sapiens
FEATURE:	
NAME/KEY:	SITE
LOCATION:	(1295)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1296)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1297)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1298)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1299)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1300)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1301)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1302)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1303)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1304)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1305)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1306)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1307)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1308)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1309)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1310)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1311)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1312)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1313)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1314)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1315)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1316)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1317)

```

1 NAME/KEY: SITE
2 LOCATION: (1342)
3 OTHER INFORMATION: n equals a,t,g, or c
4 NAME/KEY: SITE
5 LOCATION: (1343)
6 OTHER INFORMATION: n equals a,t,g, or c
7 NAME/KEY: SITE
8 LOCATION: (1344)
9 OTHER INFORMATION: n equals a,t,g, or c
10 NAME/KEY: SITE
11 LOCATION: (1345)
12 OTHER INFORMATION: n equals a,t,g, or c
13 NAME/KEY: SITE
14 LOCATION: (1346)
15 OTHER INFORMATION: n equals a,t,g, or c
16 NAME/KEY: SITE
17 LOCATION: (1347)
18 OTHER INFORMATION: n equals a,t,g, or c
19 NAME/KEY: SITE
20 LOCATION: (1348)
21 OTHER INFORMATION: n equals a,t,g, or c
22 NAME/KEY: SITE
23 LOCATION: (1349)
24 OTHER INFORMATION: n equals a,t,g, or c
25 NAME/KEY: SITE
26 LOCATION: (1350)
27 OTHER INFORMATION: n equals a,t,g, or c
28 NAME/KEY: SITE
29 LOCATION: (1351)
30 OTHER INFORMATION: n equals a,t,g, or c
31 NAME/KEY: SITE
32 LOCATION: (1352)
33 OTHER INFORMATION: n equals a,t,g, or c
34 NAME/KEY: SITE
35 LOCATION: (1353)
36 OTHER INFORMATION: n equals a,t,g, or c
37 NAME/KEY: SITE
38 LOCATION: (1354)
39 OTHER INFORMATION: n equals a,t,g, or c
40 NAME/KEY: SITE
41 LOCATION: (1355)
42 OTHER INFORMATION: n equals a,t,g, or c
43 NAME/KEY: SITE

```

Query Match	31.2%	Score 708.6	DB 11	Length 32174
Best Local Similarity	99.4%	Pred. No. 7.2e-209		
Matches 711	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Qy 1546	TTTCCCTGGTGGCCGTGTGCGGGAGCTGGGCGAAGAACTACTCGTCTCCCTTTTACAGCAC	1605		
Db 26125	TCCTTGACAGTGGCTGTGCGGGAGCTGGGCGAAGAACTACTCGTCTCCCTTTTACAGCAC	26188		
Qy 1606	TCCTTTGGCCCCAGAGCAGAGATATGAAAGCCAGGAGGTGGAAATCGATGCTTCACG	1665		
Db 26185	TCCTTTGGCCCCAGAGCAGAGATATGAAAGCCAGGAGGTGGAAATCGATGCTTCACG	26244		
Qy 1666	TGTGCTCTGTGTGCGCAGCCAGTCTTCATTGTGGGGCCAAAGGGGAAACTTTTTTTGGAG	1725		
Db 26245	TGTGCTCTGTGTGCGCAGCCAGTCTTCATTGTGGGGCCAAAGGGGAAACTTTTTTTGGAG	26304		
Qy 1726	AAGCGCTCTTGCTTTGTTCACCCACGCTGAGATGCAGTGGCGGGATCTCAAGCTCACCGCA	1785		
Db 26305	AAGCGCTCTTGCTTTGTTCACCCACGCTGAGATGCAGTGGCGGGATCTCAAGCTCACCGCA	26366		
Qy 1786	CCTCCACCTCCTGGGATTCAGATGATTTTCTCGCTCAGCCTCCCAAGTAGCTGGGAATAC	1845		
Db 26365	CCTCCACCTCCTGGGATTCAGATGATTTTCTCGCTCAGCCTCCCAAGTAGCTGGGAATAC	26422		
Qy 1846	AGGCAAGCCACCATGCCACGCTAAATTTTGTATTTTCAGTGAAAGCGGAATTTCCACAG	1905		
Db 26425	AGGCAAGCCACCATGCCACGCTAAATTTTGTATTTTCAGTGAAAGCGGAATTTCCACAG	26488		
Qy 1906	TTTGCCAGAGCTGTGTGCAATCCTTGACCGCAGATATCCACCCGCTCCGCTCCCAAA	1965		

Db	26485	TTGGCCAGGCTGGTCTCGAACTCTGACCGCAAGTATCCACCCGGCTCGGCTCCAAA	26544
Qy	1966	GTGTGGGATTTACAGCGCTGAGCCACCCTGCCCCGAAAGGGGAACTCTTTGGGAG	2025
Db	26545	GTGTGGGATTTACAGCGCTGAGCCACCGGCCCGGCCCCAAAGGGGAACTCTTTGGGAG	26604
Qy	2026	GAGCAGAGGGGCTCACAATCTCCCTCTGATTTCCCCATGCACATTTGCCCTTATCTCTCCCC	2085
Db	26605	GAGCAGAGGGGCTCACAATCTCCCTCTGATTTCCCCATGCACATTTGCCCTTATCTCTCCCC	26666
Qy	2086	ATCTAGCCAGGAATCTATTGTGTTTTCTCTTCCCAATTTACTATGATTTGTGATGTGCC	2145
Db	26665	ATCTAGCCAGGAATCTATTGTGTTTTCTCTTCCCAATTTACTATGATTTGTGATGTGCC	26724
Qy	2146	GCTACCAACCAACCCCCCATAGGGGGGGGTGAGAGGGGGTCCAAAGGCTTCCCTGCTCCACT	2205
Db	26725	GCTACCAACCAACCCCCCATAGGGGGGGGTGAGAGGGGGTCCAAAGGCTTCCCTGCTCCACT	26784
Qy	2206	TTTTCTACCTTGGACGTGATTTGATTAATCACTCTGTGTTGTTCACTTTTCA	2260
Db	26785	TTTTCTACCTTGGACGTGATTTGATTAATCACTCTGTGTTGTTCACTTTTCA	26839

```

RESULT 15
US-09-764-891-10179/c
; Sequence 10179, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10179
; LENGTH: 32174
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (29356)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29357)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29358)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29359)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29360)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29361)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29362)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29363)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29364)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE
; LOCATION: (29365)
; OTHER INFORMATION: n equals a,t,g, or c
;
; NAME/KEY: SITE

```

1	LOCATION: (29366)	n equals a,t,g, or c
2	OTHER INFORMATION: n equals a,t,g, or c	
3	NAME/KEY: SITE	
4	LOCATION: (29367)	
5	OTHER INFORMATION: n equals a,t,g, or c	
6	NAME/KEY: SITE	
7	LOCATION: (29368)	
8	OTHER INFORMATION: n equals a,t,g, or c	
9	NAME/KEY: SITE	
10	LOCATION: (29369)	
11	OTHER INFORMATION: n equals a,t,g, or c	
12	NAME/KEY: SITE	
13	LOCATION: (29370)	
14	OTHER INFORMATION: n equals a,t,g, or c	
15	NAME/KEY: SITE	
16	LOCATION: (29371)	
17	OTHER INFORMATION: n equals a,t,g, or c	
18	NAME/KEY: SITE	
19	LOCATION: (29372)	
20	OTHER INFORMATION: n equals a,t,g, or c	
21	NAME/KEY: SITE	
22	LOCATION: (29373)	
23	OTHER INFORMATION: n equals a,t,g, or c	
24	NAME/KEY: SITE	
25	LOCATION: (29374)	
26	OTHER INFORMATION: n equals a,t,g, or c	
27	NAME/KEY: SITE	
28	LOCATION: (29375)	
29	OTHER INFORMATION: n equals a,t,g, or c	
30	NAME/KEY: SITE	
31	LOCATION: (29376)	
32	OTHER INFORMATION: n equals a,t,g, or c	
33	NAME/KEY: SITE	
34	LOCATION: (29377)	
35	OTHER INFORMATION: n equals a,t,g, or c	
36	NAME/KEY: SITE	
37	LOCATION: (29378)	
38	OTHER INFORMATION: n equals a,t,g, or c	
39	NAME/KEY: SITE	
40	LOCATION: (29379)	
41	OTHER INFORMATION: n equals a,t,g, or c	
42	NAME/KEY: SITE	
43	LOCATION: (29380)	
44	OTHER INFORMATION: n equals a,t,g, or c	
45	NAME/KEY: SITE	
46	LOCATION: (29381)	
47	OTHER INFORMATION: n equals a,t,g, or c	
48	NAME/KEY: SITE	
49	LOCATION: (29382)	
50	OTHER INFORMATION: n equals a,t,g, or c	
51	NAME/KEY: SITE	
52	LOCATION: (29383)	
53	OTHER INFORMATION: n equals a,t,g, or c	
54	NAME/KEY: SITE	
55	LOCATION: (29384)	
56	OTHER INFORMATION: n equals a,t,g, or c	
57	NAME/KEY: SITE	
58	LOCATION: (29385)	
59	OTHER INFORMATION: n equals a,t,g, or c	
60	NAME/KEY: SITE	
61	LOCATION: (29386)	
62	OTHER INFORMATION: n equals a,t,g, or c	
63	NAME/KEY: SITE	
64	LOCATION: (29387)	
65	OTHER INFORMATION: n equals a,t,g, or c	
66	NAME/KEY: SITE	
67	LOCATION: (29388)	
68	OTHER INFORMATION: n equals a,t,g, or c	
69	NAME/KEY: SITE	
70	LOCATION: (29389)	
71	OTHER INFORMATION: n equals a,t,g, or c	
72	NAME/KEY: SITE	
73	LOCATION: (29390)	

1	OTHER INFORMATION:	n	equals	a,t,g,	or c
2	NAME/KEY: SITE				
3	LOCATION: (29391)				
4	OTHER INFORMATION:	n	equals	a,t,g,	or c
5	NAME/KEY: SITE				
6	LOCATION: (29392)				
7	OTHER INFORMATION:	n	equals	a,t,g,	or c
8	NAME/KEY: SITE				
9	LOCATION: (29393)				
10	OTHER INFORMATION:	n	equals	a,t,g,	or c
11	NAME/KEY: SITE				
12	LOCATION: (29394)				
13	OTHER INFORMATION:	n	equals	a,t,g,	or c
14	NAME/KEY: SITE				
15	LOCATION: (29395)				
16	OTHER INFORMATION:	n	equals	a,t,g,	or c
17	NAME/KEY: SITE				
18	LOCATION: (29396)				
19	OTHER INFORMATION:	n	equals	a,t,g,	or c
20	NAME/KEY: SITE				
21	LOCATION: (29397)				
22	OTHER INFORMATION:	n	equals	a,t,g,	or c
23	NAME/KEY: SITE				
24	LOCATION: (29398)				
25	OTHER INFORMATION:	n	equals	a,t,g,	or c
26	NAME/KEY: SITE				
27	LOCATION: (29399)				
28	OTHER INFORMATION:	n	equals	a,t,g,	or c
29	NAME/KEY: SITE				
30	LOCATION: (29400)				
31	OTHER INFORMATION:	n	equals	a,t,g,	or c
32	NAME/KEY: SITE				
33	LOCATION: (29401)				
34	OTHER INFORMATION:	n	equals	a,t,g,	or c
35	NAME/KEY: SITE				
36	LOCATION: (29402)				
37	OTHER INFORMATION:	n	equals	a,t,g,	or c
38	NAME/KEY: SITE				
39	LOCATION: (29403)				
40	OTHER INFORMATION:	n	equals	a,t,g,	or c
41	NAME/KEY: SITE				
42	LOCATION: (29404)				
43	OTHER INFORMATION:	n	equals	a,t,g,	or c
44	NAME/KEY: SITE				
45	LOCATION: (29405)				
46	OTHER INFORMATION:	n	equals	a,t,g,	or c
47	NAME/KEY: SITE				
48	LOCATION: (29406)				
49	OTHER INFORMATION:	n	equals	a,t,g,	or c
50	NAME/KEY: SITE				
51	LOCATION: (29407)				
52	OTHER INFORMATION:	n	equals	a,t,g,	or c
53	NAME/KEY: SITE				
54	LOCATION: (29408)				
55	OTHER INFORMATION:	n	equals	a,t,g,	or c
56	NAME/KEY: SITE				
57	LOCATION: (29409)				
58	OTHER INFORMATION:	n	equals	a,t,g,	or c
59	NAME/KEY: SITE				
60	LOCATION: (29410)				
61	OTHER INFORMATION:	n	equals	a,t,g,	or c
62	NAME/KEY: SITE				
63	LOCATION: (29411)				
64	OTHER INFORMATION:	n	equals	a,t,g,	or c
65	NAME/KEY: SITE				
66	LOCATION: (29412)				
67	OTHER INFORMATION:	n	equals	a,t,g,	or c
68	NAME/KEY: SITE				
69	LOCATION: (29413)				
70	OTHER INFORMATION:	n	equals	a,t,g,	or c
71	NAME/KEY: SITE				
72	LOCATION: (29414)				
73	OTHER INFORMATION:	n	equals	a,t,g,	or c

NAME/KEY: SITE
LOCATION: (29415)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29416)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE

Query Match 31.2%; Score 708.6; DB 11; Length 32174;
Best Local Similarity 99.4%; Pred. No. 7.2e-209;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1546 TTTCCCTGTGGCTGTGGGAGCTGTGAGAACTACTGCTCTCCCTTTTACAGCAC 1605
DB 6050 TCTTGGAGGTGCTGTGGGAGCTGTGAGAACTACTGCTCTCCCTTTTACAGCAC 5991
QY 1606 TCCCTTGCCCGCAGAGCAGAGAAATGAAAAGCCAGGAGGTGAGATCGATGCTTCAGC 1665
DB 5990 TCCCTTGCCCGCAGAGCAGAGAAATGAAAAGCCAGGAGGTGAGATCGATGCTTCAGC 5931
QY 1666 TGTGCTCTGTGCTCCAGCCAAAGTCTTCAATTGGGGCCAAAGGGGAAACTTTTGTGGAG 1725
DB 5930 TGTGCTCTGTGCTCCAGCCAAAGTCTTCAATTGGGGCCAAAGGGGAAACTTTTGTGGAG 5871
QY 1726 AAGGCGCTGTGCTTGTGACCCAGCTGGAATGACAGTGGGGGATCTCAGCTCACCGCAA 1785
DB 5870 AAGGCGCTGTGCTTGTGACCCAGCTGGAATGACAGTGGGGGATCTCAGCTCACCGCAA 5811
QY 1786 CCTCCACCTGCTGGGTTCAAGTATTTCTGCTCAGCTCCCAAGTAGCTGGAAATAC 1845
DB 5810 CCTCCACCTGCTGGGTTCAAGTATTTCTGCTCAGCTCCCAAGTAGCTGGAAATAC 5751
QY 1846 AGGACGCGCCACCAATGCCCACTAATTTTGTATTTTCAGTAGAAAGGGATTTCACACG 1905
DB 5750 AGGACGCGCCACCAATGCCCACTAATTTTGTATTTTCAGTAGAAAGGGATTTCACACG 5691
QY 1906 TTGGCAGAGCTGTGCTGGAATCTCTGACCGCAAGTATTCACCGCTCCGCCCAAA 1965
DB 5690 TTGGCAGAGCTGTGCTGGAATCTCTGACCGCAAGTATTCACCGCTCCGCCCAAA 5631
QY 1966 GTGCTGGGATTACAGGCGTGAAGCCGCTCCGCGCCCAAGGGGAAACTTTGTGGAG 2025
DB 5630 GTGCTGGGATTACAGGCGTGAAGCCGCTCCGCGCCCAAGGGGAAACTTTGTGGAG 5571
QY 2026 GAGCAGAGGGGCTCACTCTCCCTGTGATTTCCCAATGCAATTGCCTTATCTCTCCC 2085
DB 5570 GAGCAGAGGGGCTCACTCTCCCTGTGATTTCCCAATGCAATTGCCTTATCTCTCCC 5511
QY 2086 ATCTAGCAGGAACTATTTGTTTCTCTGCAATTACTATGATGTGATGTGCC 2145
DB 5510 ATCTAGCAGGAACTATTTGTTTCTCTGCAATTACTATGATGTGATGTGCC 5451
QY 2146 GCTACACACACCCCGCATGGGGGGGTGAGAGGGGTGCAGAGCCCTGCTCTCCACT 2205
DB 5450 GCTACACACACCCCGCATGGGGGGGTGAGAGGGGTGCAGAGCCCTGCTCTCCACT 5391
QY 2206 TTTTCTACTTGAAGTGTATTAATAATCACTTGTGTTGTCAGTTTTTCA 2260
DB 5390 TTTTCTACTTGAAGTGTATTAATAATCACTTGTGTTGTCAGTTTTTCA 5336

Search completed: November 22, 2003, 08:15:09
Job time: 707 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comogen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2003, 02:28:33 ; Search time 135 Seconds

(without alignments)
7431.584 Million cell updates/sec

Title: US-09-938-803-26

Perfect score: 2273

Sequence: 1 959999tgaagcagatcgtt.....ttttcaaaaaaaaaaaaaa 2273

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:*

1: /cgn2_6/pdata/1/ina/5A.COMB.seq:*
2: /cgn2_6/pdata/1/ina/5B.COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A.COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B.COMB.seq:*
5: /cgn2_6/pdata/1/ina/PCUS.COMB.seq:*
6: /cgn2_6/pdata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	223.4	9.8	1043	US-09-165-868-4	Sequence 4, Appli
C 2	223.4	9.8	1601	US-09-016-434-1218	Sequence 1218, Ap
C 3	221.8	9.8	90541	US-09-759-359A-3	Sequence 3, Appli
C 4	221.4	9.7	15977	US-09-608-285A-59	Sequence 59, Appli
C 5	218.8	9.6	14747	US-09-608-285A-42	Sequence 42, Appli
C 6	218.6	9.6	14747	US-09-557-800C-42	Sequence 42, Appli
C 7	218.6	9.6	62804	US-09-800-960-3	Sequence 3, Appli
C 8	218.4	9.6	283	US-08-579-445-26	Sequence 26, Appli
C 9	218	9.6	282	US-08-133-629-8	Sequence 8, Appli
C 10	217.8	9.6	14581	US-08-520-372D-4	Sequence 4, Appli
C 11	217.8	9.6	22481	US-08-367-841A-43	Sequence 43, Appli
C 12	217.8	9.6	22481	PCT-US95-07201-43	Sequence 43, Appli
C 13	217.8	9.6	22484	US-09-875-223-2	Sequence 2, Appli
C 14	215.6	9.5	21784	US-09-820-002-3	Sequence 3, Appli
C 15	215.6	9.5	42571	US-09-810-347-3	Sequence 3, Appli
C 16	215.4	9.5	174493	US-09-804-471A-3	Sequence 3, Appli
C 17	215	9.5	2932	US-09-016-434-1419	Sequence 1419, Ap
C 18	214.6	9.4	20303	US-08-370-975B-6	Sequence 6, Appli
C 19	214.6	9.4	26764	US-08-370-975B-1	Sequence 1, Appli
C 20	214.4	9.4	17327	US-07-906-871-15	Sequence 15, Appli
C 21	213.8	9.4	11531	US-08-068-945A-1	Sequence 1, Appli
C 22	213.8	9.4	11531	US-08-442-806-1	Sequence 1, Appli
C 23	213.8	9.4	11531	US-09-355-295B-1	Sequence 1, Appli
C 24	213.6	9.4	43950	US-09-735-934A-3	Sequence 3, Appli
C 25	213.6	9.4	43950	US-10-060-332-3	Sequence 3, Appli
C 26	213	9.4	53332	US-09-801-861-3	Sequence 3, Appli
C 27	213	9.4	87350	US-08-781-891-79	Sequence 79, Appli

C 28	213	9.4	87350	US-09-618-166-79	Sequence 79, Appli
C 29	213	9.4	87543	US-09-791-211-3	Sequence 3, Appli
C 30	212.8	9.4	36741	US-09-301-665-3	Sequence 3, Appli
C 31	212.6	9.4	72604	US-09-268-992-7	Sequence 7, Appli
C 32	212.6	9.4	72604	US-09-657-474-7	Sequence 7, Appli
C 33	212.4	9.3	2758	US-09-620-312D-884	Sequence 884, App
C 34	212.2	9.3	55827	US-09-813-133A-3	Sequence 3, Appli
C 35	212.2	9.3	111282	US-09-754-250-3	Sequence 3, Appli
C 36	212	9.3	1988	US-08-257-963B-11	Sequence 11, Appli
C 37	212	9.3	1988	US-08-367-841A-11	Sequence 11, Appli
C 38	212	9.3	1988	PCT-US95-07201-11	Sequence 11, Appli
C 39	212	9.3	2000	US-09-705-267A-19	Sequence 19, Appli
C 40	212	9.3	5262	US-08-520-373D-5	Sequence 5, Appli
C 41	212	9.3	80246	US-09-078-294-4	Sequence 4, Appli
C 42	212	9.3	80595	US-09-078-294-4	Sequence 3, Appli
C 43	211.4	9.3	631	US-09-385-982-354	Sequence 354, App
C 44	211.4	9.3	66804	US-09-740-041-3	Sequence 3, Appli
C 45	211.2	9.3	50000	US-09-146-053-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-165-868-4/c
Sequence 4, Application US/09165868
Patent No. 6465176
GENERAL INFORMATION:
APPLICANT: Message Pharmaceuticals, Inc.
TITLE OF INVENTION: METHOD FOR IDENTIFYING COMPOUNDS
FILE REFERENCE: 50093/003W01
CURRENT FILING DATE: 1999-10-02
PCT APPLICATION NUMBER: US/09/165,868
PCT FILING DATE: 1999-09-16
SOFTWARE: FastSeq for Windows Version 3.0
SEQUENCE OF SEQ ID NOS: 11
SEQ ID NO 4
LENGTH: 1043
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-165-868-4

Query Match 9.8%; Score 223.4; DB 4; Length 1043;
Best Local Similarity 87.1%; Pred. No. 1.2e-51;
Matches 257; Mismatches 36; Indels 2; Gaps 1;

QY	1774	TTTTTTTGGAGAGGCGCTTGTCTTTCACCCGCGTGAATGACATGGCGGATCTC	1773
DB	906	TATTTTGGAGACAGATCTTGTCTTCACCCGCGTGAATGACAGGCGATATATC	847
QY	1774	AGCTCCGCAACCTCCACCTCTCGGTTCAAGTATTTCTGCTCAGCTCCCAAGT	1833
DB	846	AGCTACGCAACTTCATCTCTGTTCAAGCAATCTCTGCTCAGCTCCCAAGT	787
QY	1834	AGCTGGAAATACAG-GCAGCGCACCATGCCCCAGTAAATTTTGTATTTTCAGTAAAC	1891
DB	786	AGCTGGAAATACAGTGGCGCCGACCATGCCCCAGTAAATTTTGTATTTTGTAGTAAAC	727
QY	1892	GGGATTACACACGTTGGCGAGCGTGTCTGAATCTCTGACCGGAATGATCCACCGC	1951
DB	726	GGGATTACACACGTTGGCGAGCGTGTAGAACTCTGACCTTAAGTATTCACCGC	667
QY	1952	CTCCGCTCCCAAGTGTGGATTCACAGGCGTGAACCAACCGTGCCTCCGCAAA	2006
DB	666	CTCAGCTCCCAAGTGTGGATTCACAGGCGTGAACCAACCGGCTCCGCTAGA	612

RESULT 2
US-09-016-434-1218/c

```
; Sequence 1218, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; FILING DATE: HEREMITH
; APPLICATION NUMBER: US/09/016,434
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 845-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1218:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g186270
; US-09-016-434-1218
```

```
Query Match          9.8%; Score 223.4; DB 4; Length 1601;
Best Local Similarity 87.1%; Pred. No. 1.5e-51;
Matches 257; Conservative 0; Mismatches 36; Indels 2; Gaps 1;
```

```
QY 1714 TTTTGTGAGAGGCGTCTTGTGTCACCCAGCGTGAATGACGTGGCGGATCTC 1773
    |||||
DB 1464 TATTTTGTGAGAGAGTCTTGTCTGTCTACCCAGCGTGAATGACGTGGCGGATCTC 1405
    |||||
QY 1774 AGCTACCGGCAACTCTCCAGCTCTGGGTTCAGATTTTCCGCGCTCAGCTCCCAAGT 1833
    |||||
DB 1404 AGCTACCTGCAACTCTCCAGCTCTGGGTTCAGATTTTCCGCGCTCAGCTCCCAAGT 1345
    |||||
QY 1834 AGCTGGAATACAG--GCAGCCACATGCGCCAGCTAATTTTGTATTTTTCAGTAGAAC 1891
    |||||
DB 1344 AGCTGGAATACAGTGTGGCGCCACATGCGCGCTAATTTTGTATTTTTCAGTAGAAC 1285
    |||||
QY 1892 GGGATTTACACGCTTGGCCAGCGTGTCTTGAATCTCTGACCGCAAGTATCCACCGCG 1951
    |||||
DB 1284 GGGGTTTACACATTTGACCGAGCTGTGAGAACTCTGACCTCAAGTATCCACCGCGC 1225
    |||||
QY 1952 CTCGGCTCCCAAGTGTGGGATTACAGGCGTGAAGCAAGCGCGCGGCCCAAA 2006
    |||||
DB 1224 CTCAGCTTCCCAAGTGTGGGATTACAGGCGTGAAGCAAGCGCGCGGCCCTAGA 1170
    |||||
```

RESULT 3
US-09-759-359A-3

```
; Sequence 3, Application US/09759359A
; Patent No. 6492153
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDER, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/09/759,359A
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 90541
; TYPE: DNA
; ORGANISM: Human
; US-09-759-359A-3
```

```
Query Match          9.8%; Score 221.8; DB 4; Length 90541;
Best Local Similarity 85.5%; Pred. No. 3.1e-50;
Matches 247; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 1714 TTTTGTGAGAGGCGTCTTGTGTCACCCAGCGTGAATGACGTGGCGGATCTC 1773
    |||||
DB 80683 TTTTGTGAGATGAGTCTTGTCTGTGCGCCAGCTGAGTGGCGCATCTT 80742
    |||||
QY 1774 AGCTACCGGCAACTCTCCAGCTCTGGGTTCAGATTTTCTGCTCAGCTCCCAAGT 1833
    |||||
DB 80743 AGCTACCTGCAACTCTGCTGCTCCAGGTTCAAGCAATCTCTGCGCGCCAGCTCCCATGT 80802
    |||||
QY 1834 AGCTGGAATACAGGAGCCACCATGCGCCAGCTAATTTTGTATTTTTCAGTAGAACG 1893
    |||||
DB 80803 AGCTGGAATACAGGAGCCACCATGCGCCAGCTAATTTTGTATTTTTCAGTAGAACG 80862
    |||||
QY 1894 GATTCACACGTTGGCCAGCGTGTCTGCAATCTCTGACCGCAAGTATCCACCGCGCT 1953
    |||||
DB 80863 GATTCACGTTGTGGCCAGCGTGTCTGCAATCTCTGACCGCAAGTATCCACCGCGCT 80922
    |||||
QY 1954 CCGCTTCCCAAGTGTGGGATTACAGGCGTGAAGCAAGCGTGGCGGCC 2002
    |||||
DB 80923 TAGCTTCCCAAGTGTGGGATTACAGGCGTGAAGCAAGCGTGGCGGCC 80971
    |||||
```

```
RESULT 4
US-09-608-285A-59
; Sequence 59, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
```



```

; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 15977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: CD39-L4/L6 Gene Sequence
; NAME/KEY: CDS
; LOCATION: (245)..(461)
; NAME/KEY: CDS
; LOCATION: (1454)..(1533)
; NAME/KEY: CDS
; LOCATION: (2734)..(2877)
; NAME/KEY: CDS
; LOCATION: (4364)..(4439)
; NAME/KEY: CDS
; LOCATION: (4679)..(4714)
; NAME/KEY: CDS
; LOCATION: (5326)..(5414)
; NAME/KEY: CDS
; LOCATION: (5723)..(5802)
; NAME/KEY: CDS
; LOCATION: (6751)..(6812)
; NAME/KEY: CDS
; LOCATION: (7758)..(7859)
; NAME/KEY: CDS
; LOCATION: (8712)..(8852)
; NAME/KEY: CDS
; LOCATION: (9831)..(9887)
; NAME/KEY: CDS
; LOCATION: (11613)..(11728)
; NAME/KEY: CDS
; LOCATION: (13146)..(13691)
; NAME/KEY: CDS
; LOCATION: (15702)..(15839)
; NAME/KEY: misc feature
; LOCATION: (14871)
; OTHER INFORMATION: n = a or c or g or t
; US-09-608-285A-59

Query Match          9.7%; Score 221.4; DB 4; Length 15977;
Best Local Similarity 87.0%; Pred. No. 1.7e-50;
Matches 253; Conservative 0; Mismatches 36; Indels 2; Gaps 1;

QY 1713 CTTTCTTTTGGAGAGGCGCTCTTGTCTTGTCAACCGCTGGATGTCAGTGCGGGATCT 1772
DB 11237 CTTTCTTTTGGAGAGAGTCTTGTCTGTCTGTCTCAACCGCTGGATGTCAGTGCGGGATCT 11296
QY 1773 CAGCTACCGCAACTCTCACTCTGGGTTCAAGTATTTTCTGCTCAAGCTTCCAG 1832
DB 11297 GGGCTACGTCAACTCTGCTCTTCCAGGTTCAAGGATCTCTGCTCAAGCTTCCAG 11356
QY 1893 TAGCTGGGAATACA--GGCAGCGCCACATGCCCACTAATTTTCTATTTTCAAGTAA 1890
DB 11337 TAGCTGGGAATACAAGGATGTCAGTCAACGCGCTCTAATTTTCTATTTTCAAGTAA 11416
QY 1891 GGGATTTACCACTGTCAGGCTGTCGATCTGCAACTCTGCAACCGCAAGTATCCACCG 1950
DB 11417 TGGGCTTTCACCACTGTCAGGCTGTCGATCTGCAACTCTGCAACCGCAAGTATCCACCG 11476
QY 1951 CTTCCGCTCCCAAGTGTGGGATTTACAGGCTGACCCAGCCGCGCGCC 2003
DB 11477 CTTCCGCTCCCAAGTGTGGGATTTACAGGCTGACCCAGCCGCGCGCC 11529

RESULT 5
US-09-608-285A-42
; Sequence 42, Application US/09608285A
; Patent No. 6335013
```

```

; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 14747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13641)
; OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
; US-09-608-285A-42

Query Match          9.6%; Score 218.8; DB 4; Length 14747;
Best Local Similarity 86.6%; Pred. No. 8.3e-50;
Matches 253; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1714 TTTTCTTTTGGAGAGGCGCTCTTGTCTTGTCAACCGCTGGATGTCAGTGCGGGATCTC 1773
DB 10008 TTTTCTTTTGGAGAGAGTCTTGTCTGTCTGTCAACCGCTGGATGTCAGTGCGGGATCTC 10067
QY 1774 AGCTACCGCAACTCTCACTCTGGGTTCAAGTATTTTCTGCTCAAGCTTCCAGT 1833
DB 10068 GGGCTACGTCAACTCTGCTCTTCCAGGTTCAAGGATCTCTGCTCAAGCTTCCAGT 10127
QY 1834 AGCTGGGAATACA--GGCAGCGCCACATGCCCACTAATTTTCTATTTTCAAGTAA 1891
DB 10128 AGCTGGGAATACAAGGATGTCAGTCAACGCGCTCTAATTTTCTATTTTCAAGTAA 10187
QY 1892 GGGATTTACCACTGTCAGGCTGTCGATCTGCAACTCTGCAACCGCAAGTATCCACCG 1951
DB 10188 GGGGTTTTCACCACTGTCAGGCTGTCGATCTGCAACTCTGCAACCGCAAGTATCCACCG 10247
QY 1952 CTTCCGCTCCCAAGTGTGGGATTTACAGGCTGACCCAGCCGCGCGCC 2003
DB 10248 CTTCCGCTCCCAAGTGTGGGATTTACAGGCTGACCCAGCCGCGCGCC 10299

RESULT 6
US-09-557-800C-42
; Sequence 42, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
```

APPLICANT: Young, George
TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
FILE REFERENCE: 28110/36457
CURRENT APPLICATION NUMBER: US/09/557,800C
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/122449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/244444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/118,205
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42
LENGTH: 14747
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (13641)
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-557-800C-42

Query Match 9.6%; Score 218.8; DB 4; Length 14747;
Best Local Similarity 86.6%; Pred. No. 8.3e-50;
Matches 253; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1714 TTTTGGAGAGGCGTCTTGGTTTGTACCCAGCTGGAATGAGTGGCGGATCTC 1773
DB 10008 TTTTGGAGAGGCGTCTTGGTTTGTACCCAGCTGGAATGAGTGGCGGATCTC 10067
QY 1774 AGCTACCCGACCTCCAGCTCGGTTCAAGTATTTCCGCGCTCAGCTCCGCAAGT 1833
DB 10068 GGGTCTGACACCTCTGCTTCCAGTTCAAGGATTTCTGCTCAGCTCCGCAAGT 10127
QY 1834 AGCTGGGAATACA--GGCAGCCGACCATGCGCAGCTAATTTTGTATTTTCACTAGAAAC 1891
DB 10128 AGCTGGGAATACAAGCGGTACACCAACGCGGTCTAATTTTGTATTTTGTAGAGAT 10187
QY 1892 GGGATTTACCAAGTTGGCAGGCTGTCTGAACTCTGACCGCAAGTATCCACCGC 1951
DB 10188 GGGGTTTACCAAGTTGGCAGGCTGTCTGAACTCTGACCGCAAGTATCCACCGC 10247
QY 1952 CTCGCGCTCCCAAGTCTGGATTACAGGCGTGAGCCACGCGCCGCC 2003
DB 10248 CTCAGCTCCCAAGTCTGGATTACAGGCGTGAGCCACGCGCTGCC 10299

RESULT 7
US-09-800-960-3/C
Sequence 3, Application US/09800960
Patent No. 6387677
GENERAL INFORMATION:
APPLICANT: YE, Jane et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL001158
CURRENT APPLICATION NUMBER: US/09/800,960
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FaestSeq for Windows Version 4.0

SEQ ID NO 3
LENGTH: 62804
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1) (62804)
OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

Query Match 9.6%; Score 218.6; DB 4; Length 62804;
Best Local Similarity 86.1%; Pred. No. 1.9e-49;
Matches 254; Conservative 0; Mismatches 39; Indels 2; Gaps 1;

QY 1714 TTTTGGAGAGGCGTCTTGGTTTGTACCCAGCTGGAATGAGTGGCGGATCTC 1773
DB 13733 TTTCTTGTGAGAGCGGAGTCTGTTGTCAGGCTGGGGTGAGTGTGATCTC 13674
QY 1774 AGCTACCCGACCTCCAGCTCGGTTCAAGTATTTTCTGCTCAGCTCCGCAAGT 1833
DB 13673 AGCTACCTCAACCTTGCTCTGCTGTTCAAGCAATTCCTGCTCAGCTCCGCAAGT 13614
QY 1834 AGCTGGGAATACAG--CAGCCACCATGCGCAGCTAATTTTGTATTTTCACTAGAAAC 1891
DB 13613 AGCTGGGAATACAGGACACACACAGCTGCTGCTAATTTTGTATTTTGTAGAGAT 13554
QY 1892 GGGATTTACCAAGTTGGCAGGCTGTCTGAACTCTGACCGCAAGTATCCACCGC 1951
DB 13553 GGGGTTTACCAAGTTGGCAGGCTGTCTGAACTCTGACCTGAGTATCCACCGC 13494
QY 1952 CTCGCGCTCCCAAGTCTGGATTACAGGCGTGAGCCACGCGCCGCC 2006
DB 13493 CTCGCGCTCCCAAGTCTGGATTACAGGCGTGAGCCACGCGCTGCC 13439

RESULT 8
US-08-579-445-26/C

Sequence 26, Application US/08579445
Patent No. 6566053
GENERAL INFORMATION:
APPLICANT: Peruchio, Manuel
APPLICANT: Ionov, Yuri
APPLICANT: Malkhosyan, Sergei
TITLE OF INVENTION: Identification of Neoplasms by Detection
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/579,445
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/152,484
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kirkpatrick, Anita M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: STRATAG.009A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

Db 9590 AGCTCACTGCAACCTCCACTCTCTGATTCAACTGATTCCTCCGCTTACCTCCCAAGT 9649
QY 1834 AGCTGGAAATACAGGCA--CGCCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAG 1891
Db 9650 AGCTGGAGATTAACAGGATGCGCCACACCCGCTAATTTTGTATTTTGTAGTAGAGT 9709
QY 1892 GGGATTTCACCAAGTGGCGGAGCTGTCTCGAATCTCGACCGCAAGTAGATCCACCCGC 1951
Db 9710 GGGGTTTCTCAGTGTGGCCAGGATGCTCAAACTCTGACCTCAGGTATCTACCCGC 9769
QY 1952 CTCGGCTCCCAAGTGTGGATTACAGGCGTAGCCACCGTCCCGGCC 2002
Db 9770 CTCGGCTCTCAAGTGTGGATTACAGGTTTGAGCACTGCGCTCGCC 9820

RESULT 11
US-08-367-841A-43

; Sequence 43, Application US/08367841A

; Patent No. 6319687

; GENERAL INFORMATION:

; APPLICANT: Chader, Gerald J.; Rodriguez,
; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
; APPLICANT: Tombrian-Tink, Joyce

; TITLE OF INVENTION: PIGMENT EPITHELIUM

; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC

; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morgan & Finnegan

; STREET: 345 Park Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/367,841A

; FILING DATE: 30-DEC-1994

; CLASSIFICATION: 435

; APPLICATION NUMBER: 08/257,963

; FILING DATE: 07-JUN-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/952,796

; FILING DATE: 24-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: DOROTHY R. AUTH

; REGISTRATION NUMBER: 36434

; REFERENCE/DOCKET NUMBER: 20264126US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22481 Base Pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Double

; TOPOLOGY: Unknown

; MOLECULE TYPE: Genomic DNA

; FEATURE:

; NAME/KEY: P1-147

; LOCATION:

; IDENTIFICATION METHOD:

; OTHER INFORMATION: full length genomic

; OTHER INFORMATION: sequence for PEDF plus flanking sequences.

; US-08-367-841A-43

Query Match 9.6%; Score 217.8; DB 4; Length 22481;

Best Local Similarity 86.6%; Pred. No. 1,9e-49;
Matches 222; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1714 TTTTGTGGAAGAGCGTCTTGTCTTGTACCCACGCTGAATGAGTGGGGATCTC 1773
Db 9519 TTTTGTGGAATGAGTCTTGTCTTGTACCCACGCTGAATGAGTGGGGATCTC 9578
QY 1774 AGTCAACCGCAACCTCCACTCTCTGATTAAGTATTTCTGCTCAGCTCCCAAGT 1833
Db 9579 AGTCACTCAACCTCCACTCTCTGATTAAGTATTTCTGCTCAGCTCCCAAGT 9638
QY 1834 AGCTGGAAATACAGGCA--CGCCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAG 1891
Db 9639 AGCTGGAGATTAACAGGATGCGCCACACCCGCTAATTTTGTATTTTGTAGTAGAGT 9698
QY 1892 GGGATTTCACCAAGTGGCGGAGCTGTCTCGAATCTCGACCGCAAGTAGATCCACCCGC 1951
Db 9699 GGGGTTTCTCAGTGTGGCCAGGATGCTCAAACTCTGACCTCAGGTATCTACCCGC 9758
QY 1952 CTCGGCTCCCAAGTGTGGATTACAGGCGTAGCCACCGTCCCGGCC 2002
Db 9759 CTCGGCTCTCAAGTGTGGATTACAGGTTTGAGCACTGCGCTCGCC 9809

RESULT 12
PCT-US95-07201-43

; Sequence 43, Application PC/TUS9507201

; GENERAL INFORMATION:

; APPLICANT: Chader, Gerald J.; Becerra, Sofia

; APPLICANT: Patricia; Schwartz, Joan P.;

; APPLICANT: Taniwaki, Takayuki

; TITLE OF INVENTION: PIGMENT EPITHELIUM

; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC

; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morgan & Finnegan, L.L.P.

; STREET: 345 Park Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/07201

; FILING DATE: 06-JUN-1995

; CLASSIFICATION:

; APPLICATION NUMBER: 08/367,841

; FILING DATE: 30-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/257,963

; FILING DATE: 07-JUN-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/952,796

; FILING DATE: 24-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: DOROTHY R. AUTH

; REGISTRATION NUMBER: 36434

; REFERENCE/DOCKET NUMBER: 20264126PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22481 Base Pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Double

; TOPOLOGY: Unknown

Query Match 9.6%; Score 217.8; DB 4; Length 22481;

MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Pl-147
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: full length genomic
OTHER INFORMATION: Sequence for PEDF plus flanking sequences.
PCT-US95-07201-43

Query Match
Best Local Similarity 9.6%; Score 217.8; DB 5; Length 22484;
Matches 252; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1714 TTTTGTGAGAGAGCGCTTGTGCTTGTGACACCGTGAATGAGTGGGAGATCTC 1773
DB 9519 TTTTGTGAGATGAGTGTGCTGCTGTTGCCACCTGAGTGAAGTGGCCGATCTC 9578
QY 1774 AGCTACCGCAACCTCCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGT 1833
DB 9579 AGCTCACTGCAACCTCCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGT 9638
QY 1834 AGCTGGGATATACAGCA--CGCACATGCCCACTAATTTTGTATTTTCACTAAGAAC 1891
DB 9639 AGCTGGGATATACAGCAAGATGCGCACACCGGCTAATTTTGTATTTTGTAGTAAAGT 9698
QY 1892 GGGATTTACACAGTGGCGAGGCTGTCTGCACTCTGACCGCAAGTATCCACCGCC 1951
DB 9699 GGGGTTTCTCCAGTGTGGCGAGATGTCTCAAACTCTGACCTCAAGTATCTACCCGC 9758
QY 1952 CTCGCGCTCCCAAGTGTGCTGGATTAACAGCGCTGAGCCACCGTCCCGGCC 2002
DB 9759 CTCGCGCTCTCAAGTGTGCTGGATTAACAGCGCTGAGCCACCGTCCCGGCC 9809

RESULT 13

US-09-875-223-2
Sequence 2, Application US/09875223
Patent No. 6391850
GENERAL INFORMATION:
APPLICANT: No. 6391850thwestern University
APPLICANT: No. 63918501 Bouck
APPLICANT: David Dawson
APPLICANT: Paul Gillette
TITLE OF INVENTION: Method and Compositions for Inhibiting Angiogenesis
FILE REFERENCE: 0290-2303
CURRENT APPLICATION NUMBER: US/09/875,223
CURRENT FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 09/122,079
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: PCT/US98/15228
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: US 08/899,304
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 22484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Unsure
LOCATION: 1...22484
OTHER INFORMATION: "n" means either a, c, t, or g
US-09-875-223-2

Query Match
Best Local Similarity 9.6%; Score 217.8; DB 4; Length 22484;
Matches 252; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1714 TTTTGTGAGAGAGCGCTTGTGCTTGTGACACCGTGAATGAGTGGGAGATCTC 1773
DB 9519 TTTTGTGAGATGAGTGTGCTGCTGTTGCCACCTGAGTGAAGTGGCCGATCTC 9578

QY 1774 AGCTACCGCAACCTCCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGT 1833
DB 9579 AGCTCACTGCAACCTCCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGT 9638
QY 1834 AGCTGGGATATACAGCA--CGCACATGCCCACTAATTTTGTATTTTCACTAAGAAC 1891
DB 9639 AGCTGGGATATACAGCAAGATGCGCACACCGGCTAATTTTGTATTTTGTAGTAAAGT 9698
QY 1892 GGGATTTACACAGTGGCGAGGCTGTGCTGCACTCTGACCGCAAGTATCCACCGCC 1951
DB 9699 GGGGTTTCTCCAGTGTGGCGAGATGTCTCAAACTCTGACCTCAAGTATCTACCGCC 9758
QY 1952 CTCGCGCTCCCAAGTGTGCTGGATTAACAGCGCTGAGCCACCGTCCCGGCC 2002
DB 9759 CTCGCGCTCTCAAGTGTGCTGGATTAACAGCGCTGAGCCACCGTCCCGGCC 9809

RESULT 14

US-09-820-002-3
Sequence 3, Application US/09820002
Patent No. 6482630
GENERAL INFORMATION:
APPLICANT: Gan, Weiniu
APPLICANT: Ye, Jane
APPLICANT: DiFrancesco, Valentina
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
FILE REFERENCE: CLO01194
CURRENT APPLICATION NUMBER: US/09/820,002
CURRENT FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 21784
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(21784)
OTHER INFORMATION: n = A,T,C or G
US-09-820-002-3

Query Match
Best Local Similarity 9.5%; Score 215.6; DB 4; Length 21784;
Matches 263; Conservative 0; Mismatches 59; Indels 2; Gaps 1;

QY 1714 TTTTGTGAGAGAGCGCTTGTGCTTGTGACACCGTGAATGAGTGGGAGATCTC 1773
DB 18045 TTTTGTGAGAGAGAGTCTTGTCTGTCTGCTGCTGAGTGAATGACATATCTC 18104
QY 1774 AGCTACCGCAACCTCCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGT 1833
DB 18105 GGGTCACTGCAACCTCCACCTCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGT 18164
QY 1834 AGCTGGGATATACAG--GCAAGCCACATGCCCACTAATTTTGTATTTTCACTAAGAAC 1891
DB 18165 AGCTGGGATATACAGTGTGGTGTGACACATGCCCGCTAATTTTGTATTTTGTAGTAAAGT 18224
QY 1892 GGGATTTACACAGTGGCGAGGCTGTGCTGCACTCTGACCGCAAGTATCCACCGCC 1951
DB 18225 GAGGTTTGGCAGTGTGGCGAGGCTGTCTCAAACTCTGACATCAAGTATCTCTCTGC 18284
QY 1952 CTCGCGCTCCCAAGTGTGAGATTAACAGCGCTGAGCCACCGTCCCGGCCCAAGGGA 2011
DB 18285 CTCGCGCTCCCAAGTGTGAGATTAACAGCGCTGAGCCACCGTCCCGGCCCAAGGGA 2011
QY 2012 AACTCTGTGGAGAGAGAGAGG 2035
DB 18345 TTTTGTGAGAGAGAGAGAGAGG 18368

```

RESULT 15
US-09-810-347-3
; Sequence 3, Application US/09810347
; Patent No. 6461847
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001169
; CURRENT APPLICATION NUMBER: US/09/810,347
; CURRENT FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 42571
; TYPE: DNA
; ORGANISM: Human
US-09-810-347-3

```

```

Query Match          9.5%; Score 215.6; DB 4; Length 42571;
Best Local Similarity 86.0%; Pred. No. 1.1e-48;
Matches 251; Conservative 0; Mismatches 39; Indels 2; Gaps 1;

QY 1717 TTTTGGAGAGGCGCTTGTCTTGTCCACCCAGCTGGAATGCAGTGGCGGATCTCAGC 1776
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36537 TTTTGGAGATAGAGTCTGCTGTGTTACCCAGGCTGAGTGCAGTGGCGCAATCTCGGC 36596

QY 1777 TCACCGCAACCTCCACCTCCTGGGTTCAAGTGAATTTCTGCTCAGCCTCCCAAGTAGC 1836
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36597 TCACCTGCACCTCTCACCCTCTGTGTTCAAGAATTCCTGCTCAGCCTCTGAGTAGC 36656

QY 1837 TGGGAATACAG--GCACGCCACCATGCGCAGCTAATTTTGTATTTTTCAGTAGAAGCGG 1894
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36657 TGGGATTAACAGGCGCCGCCACCATGCGCAGCTAATTTTGTATTTTTCAGTAGAAGCGG 36716

QY 1895 ATTTCACACAGTGGCGCAGGCTGCTCGAATCCTGACCGCAAGTATCCACCGCTC 1954
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36717 GTTTCCTCCATGTGGCCAGGCTGCTCGAATCCTGACCTCAGTGAATTCACCGCTG 36776

QY 1955 CGCCTCCCAAGTGTGGATTAACAGGCGTGAAGCCACCGCGGCCAA 2006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36777 GGCCTTCCAAATGCTGGGATTAACAGGCGTGAAGCCACCATGCTGCCCAA 36828

```

Search completed: November 22, 2003, 07:54:51
 Job time : 138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using ew model

Run on: November 22, 2003, 10:43:46 ; Search time 704 Seconds
(without alignments)
10552.029 Million cell updates/sec

Title: US-09-938-803-26

Perfect score: 2273

Sequence: 1 99999tcgaagcagatcgtt.....tttccaaaaa 2273

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2169961 seqs, 1634102185 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2273	100.0	2273	9	US-09-938-803-26
2	1842	81.0	1842	10	US-09-880-107-3415
3	1842	81.0	1842	10	US-09-954-531-1368
4	708	31.1	7461	9	US-09-764-860-1135
5	708	31.1	7461	10	US-09-764-904-91
6	708	31.1	7461	14	US-10-091-548-91
7	708	31.1	7461	14	US-10-074-095-1135
8	708	31.1	32174	9	US-09-808-711-158
9	708	31.1	32174	9	US-09-764-860-1134
10	708	31.1	32174	10	US-09-864-877-2645
11	708	31.1	32174	10	US-09-860-670-232
12	708	31.1	32174	10	US-09-764-904-90
13	708	31.1	32174	11	US-09-764-891-6480
14	708	31.1	32174	11	US-09-764-891-10135
15	708	31.1	32174	11	US-09-764-891-10179
16	708	31.1	32174	14	US-10-091-548-90

C 17	708	31.1	32174	14	US-10-074-095-1134	Sequence 1134, Ap
18	428	18.8	459	11	US-09-918-995-22426	Sequence 22426, A
19	392	17.2	475	11	US-09-918-995-26257	Sequence 26257, A
20	355	15.6	363	11	US-09-918-995-30086	Sequence 30086, A
21	317	13.9	490	11	US-09-918-995-11899	Sequence 11899, A
22	254	11.2	990	10	US-09-964-8244-271	Sequence 271, App
23	254	11.2	990	10	US-09-954-531-973	Sequence 973, App
24	254	11.2	990	10	US-09-954-531-1369	Sequence 1369, App
25	181	8.0	226	12	US-10-029-386-16123	Sequence 16123, A
26	181	8.0	505	12	US-10-029-386-2423	Sequence 2423, A
27	167	7.3	173	12	US-10-029-386-16861	Sequence 16861, A
28	167	7.3	599	12	US-10-029-386-3161	Sequence 3161, Ap
29	163	7.2	478	9	US-09-864-761-87	Sequence 87, Appl
30	153	6.7	758	12	US-10-027-632-8362	Sequence 8362, Ap
31	153	6.7	758	13	US-10-027-632-8362	Sequence 8362, Ap
32	153	6.7	2006	12	US-10-027-632-99441	Sequence 99441, A
33	153	6.7	2006	13	US-10-027-632-99441	Sequence 99441, A
34	135	5.9	489	11	US-09-918-995-1933	Sequence 1933, Ap
35	128	5.6	177	9	US-09-864-761-16925	Sequence 16925, A
36	112	4.9	217	10	US-09-796-692-7696	Sequence 7696, Ap
37	112	4.9	217	14	US-10-040-862-7696	Sequence 7696, Ap
38	73	3.2	280	10	US-09-880-107-1992	Sequence 1992, Ap
39	59	2.6	74962	14	US-10-274-974-3	Sequence 3, Appl1
40	55	2.4	923	12	US-10-027-632-10590	Sequence 10590, A
41	55	2.4	923	13	US-10-027-632-10590	Sequence 10590, A
42	53	2.3	10901	11	US-09-764-891-5924	Sequence 5924, Ap
43	52	2.3	819	12	US-10-027-632-140838	Sequence 140838, A
44	52	2.3	819	13	US-10-027-632-140838	Sequence 140838, A
45	52	2.3	76798	10	US-09-880-107-3949	Sequence 3949, Ap

ALIGNMENTS

RESULT 1
US-09-938-803-26
; Sequence 26, Application US/09938803
; Patent No. US20020076762A1
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Reddy, Roopa
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Yang, Junming
; APPLICANT: Azimzal, Yalda
; TITLE OF INVENTION: FULL-LENGTH EXPRESSED GENETIC MARKERS
; FILE REFERENCE: PF-0695 US
; CURRENT APPLICATION NUMBER: US/09/938,803
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/311,894
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 2273
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte Clone 1928920
US-09-938-803-26

Query Match 100.0%; Score 2273; DB 9; Length 2273;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGTGAAGCATGTTTGGCCGATTTGGGGGGGGGGGAGTGGGGGGTCCCTGG 60
DB 1 GGGGGTGAAGCATGTTTGGCCGATTTGGGGGGGGGGGAGTGGGGGGTCCCTGG 60
QY 61 GGGCTCCGAGATTAAATGGCTTCAGCGAGAGGAGACGAGGAGACTGTGTGGCCG 120

Dh 61 GGGCTCCGGAGTTAAGATGGCGTCTCAGCGAGGGGGAGCAGAGGGACGTGTGGCGC 120
Qy 121 TGGCGGGGGTTCTGCACTCGGGTTTCCAGAGCTGAGCCTTAAACAAGTTGGCACTGCC 180
Dh 121 TGGCGGGGGTTCTGCACTCGGGTTTCCAGAGCTGAGCCTTAAACAAGTTGGCACTGCC 180
Qy 181 TGGCGGGGGTGAAGACGGGGCTGGGGTGAATCTGCACTCTCTGGGGTTTACCCCTTGG 240
Dh 181 TGGCGGGGGTGAAGACGGGGCTGGGGTGAATCTGCACTCTCTGGGGTTTACCCCTTGG 240
Qy 241 CTCTGTTTATCGGCACTTACCTTTTTCAGAGGAGACTTACCTACCTCCTCTCCATA 300
Dh 241 CTCTGTTTATCGGCACTTACCTTTTTCAGAGGAGACTTACCTCCTCCTCTCCATA 300
Qy 301 CCTTTACAGGCTCTCAATTGCTTATTTTAACTTTGAAACGAGCTTACCACTCCCTGC 360
Dh 301 CCTTTACAGGCTCTCAATTGCTTATTTTAACTTTGAAACGAGCTTACCACTCCCTGC 360
Qy 361 TGTGATATGCTTCAGTTCCTCATCTTGAGACTAATGGGCGGACCATCATGCGGTCC 420
Dh 361 TGTGATATGCTTCAGTTCCTCATCTTGAGACTAATGGGCGGACCATCATGCGGTCC 420
Qy 421 TCACTACCTTTTGGCTTCAGATGGCTTACCTTGGCTGATATTAATTAAGTCCACG 480
Dh 421 TCACTACCTTTTGGCTTCAGATGGCTTACCTTGGCTGATATTAATTAAGTCCACG 480
Qy 481 GCAACTACGATATCAAGTGAACAATGCAATGCTGTCTGACTTGAACCTGATGGTT 540
Dh 481 GCAACTACGATATCAAGTGAACAATGCAATGCTGTCTGACTTGAACCTGATGGTT 540
Qy 541 TGGCTGTTGACTACTTTGACGAGGAGGAAAGATCAGATTCCTTGTCTGAGCAACAG 600
Dh 541 TGGCTGTTGACTACTTTGACGAGGAGGAAAGATCAGATTCCTTGTCTGAGCAACAG 600
Qy 601 AATATGCAATACGTGTGTCTCTCCCTGCTGGAAGTGTGTGTTCTCTACTTCTATG 660
Dh 601 AATATGCAATACGTGTGTCTCTCCCTGCTGGAAGTGTGTGTTCTCTACTTCTATG 660
Qy 661 GGGCTCTTGTGTAGGGGCCCCAGTTCTCATGTAATCTACATGAAGCTGTGAGGAG 720
Dh 661 GGGCTCTTGTGTAGGGGCCCCAGTTCTCATGTAATCTACATGAAGCTGTGAGGAG 720
Qy 721 AGCTGATTGACATCCAGAAAGATACCAACAGCATCATTCCTGCTCTGAGGCGCTGA 780
Dh 721 AGCTGATTGACATCCAGAAAGATACCAACAGCATCATTCCTGCTCTGAGGCGCTGA 780
Qy 781 GTCTGGGCTTTTCTACCTAGTGGGTACACACTGCTCAGCCCCACATCAGAAAGACT 840
Dh 781 GTCTGGGCTTTTCTACCTAGTGGGTACACACTGCTCAGCCCCACATCAGAAAGACT 840
Qy 841 ATCTCCCTCAGTAAGATATGACAAACACCCCTTGTGGTCCGCTGATGTATCTGTA 900
Dh 841 ATCTCCCTCAGTAAGATATGACAAACACCCCTTGTGGTCCGCTGATGTATCTGTA 900
Qy 901 TCTGGGGCAAGTTGTCTGTACAAATATGTCACTGTGGTGTGACAGAAAGAGAT 960
Dh 901 TCTGGGGCAAGTTGTCTGTACAAATATGTCACTGTGGTGTGACAGAAAGAGAT 960
Qy 961 GCATTTTGAAGCGGCTGGGCTTCAATGCTTTGAAAGAAAGGGCAAGCAAGTGGATG 1020
Dh 961 GCATTTTGAAGCGGCTGGGCTTCAATGCTTTGAAAGAAAGGGCAAGCAAGTGGATG 1020
Qy 1021 CCTGTGCAACGAAAGTGTGGCTTTTGAAGAAACCCCGCTTACATGGGCACTTGG 1080
Dh 1021 CCTGTGCAACGAAAGTGTGGCTTTTGAAGAAACCCCGCTTACATGGGCACTTGG 1080
Qy 1081 CCTGATTCACATCAACCAACGCTGGGTGGGCGGCTACATTTTCAAGAGACTCACT 1140
Dh 1081 CCTGATTCACATCAACCAACGCTGGGTGGGCGGCTACATTTTCAAGAGACTCACT 1140
Qy 1141 TCTTTGAAATTAAGAACTCTCTCAGGGGTCTCTGTTGTATTTCTGGCCCTCTGGACG 1200
Dh 1141 TCTTTGAAATTAAGAACTCTCTCAGGGGTCTCTGTTGTATTTCTGGCCCTCTGGACG 1200

Qy 1201 GCTTGACCTCAGATATCTGTGTCTCTCCAGATGGAATTCCTCATTTGTTATGTGAAA 1260
Dh 1201 GCTTGACCTCAGATATCTGTGTCTCTCCAGATGGAATTCCTCATTTGTTATGTGAAA 1260
Qy 1261 GACAGGCTGACAGGCTCATTTCAAGAGGCCCCCTGAGCAAGCTGGCGGCAATACG 1320
Dh 1261 GACAGGCTGACAGGCTCATTTCAAGAGGCCCCCTGAGCAAGCTGGCGGCAATACG 1320
Qy 1321 TCTTCCAGCCCTTCTACTATTTGTGCAACAGCATCACTGGCTTTTATGAGTTACT 1380
Dh 1321 TCTTCCAGCCCTTCTACTATTTGTGCAACAGCATCACTGGCTTTTATGAGTTACT 1380
Qy 1381 CCATGATGCTTCTGCTCTTCTCAGTGGGACAAATGGCTTAAAGTGTATTAATTCATCT 1440
Dh 1381 CCATGATGCTTCTGCTCTTCTCAGTGGGACAAATGGCTTAAAGTGTATTAATTCATCT 1440
Qy 1441 ATTTCTTGGCCACATCTTCTCTGAGCCTACTATTCATATTTGCTTATTAACAAAG 1500
Dh 1441 ATTTCTTGGCCACATCTTCTCTGAGCCTACTATTCATATTTGCTTATTAACAAAG 1500
Qy 1501 CAATGCTGCCAAGAAAGAAAGTTAAAGAAATGGAATATTCATTTCTCTGGTGGCT 1560
Dh 1501 CAATGCTGCCAAGAAAGAAAGTTAAAGAAATGGAATATTCATTTCTCTGGTGGCT 1560
Qy 1561 GTGGGGGACTGTGTGCAAAATCTACTGCTCTCTTTCACAGACTCTCTTGGCCAGAG 1620
Dh 1561 GTGGGGGACTGTGTGCAAAATCTACTGCTCTCTTTCACAGACTCTCTTGGCCAGAG 1620
Qy 1621 CAGAGATGGAAGAAAGCCAGGAGGTGAAGATGATCTTCCAGCTGTGCTGTGCTGC 1680
Dh 1621 CAGAGATGGAAGAAAGCCAGGAGGTGAAGATGATCTTCCAGCTGTGCTGTGCTGC 1680
Qy 1681 AGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGGAAGAGCTGTGCTCT 1740
Dh 1681 AGCAAGTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGGAAGAGCTGTGCTCT 1740
Qy 1741 GTTACCCAGCTGGGAATGCAAGTGGCGGATCTCAGCTCACCGCACTCTCCTGGG 1800
Dh 1741 GTTACCCAGCTGGGAATGCAAGTGGCGGATCTCAGCTCACCGCACTCTCCTGGG 1800
Qy 1801 TTCAATGATTTTCTGCTCAGCTCAGCTCCCAAGATAGCTGGGAATACAGGACGCACTG 1860
Dh 1801 TTCAATGATTTTCTGCTCAGCTCAGCTCCCAAGATAGCTGGGAATACAGGACGCACTG 1860
Qy 1861 CCCAGCTAATTTTGTATTTTCAAGTAAACGGGATTTCAACAAGCTTGGCCAGCTGTCT 1920
Dh 1861 CCCAGCTAATTTTGTATTTTCAAGTAAACGGGATTTCAACAAGCTTGGCCAGCTGTCT 1920
Qy 1921 TCGAATCTGCAACCGGAATGCCACCGGCTCCGCTCCCAAGATGCTGGGATTAACG 1980
Dh 1921 TCGAATCTGCAACCGGAATGCCACCGGCTCCGCTCCCAAGATGCTGGGATTAACG 1980
Qy 1981 GCGTGAAGCAACGCTGCGCGGCCCAAGGGGAAATCTTGTGGAGAGAGAGAGGGCTCA 2040
Dh 1981 GCGTGAAGCAACGCTGCGCGGCCCAAGGGGAAATCTTGTGGAGAGAGAGAGGGCTCA 2040
Qy 2041 CATCTCCCTCTGATTTCCCATGCAACTTGCCTTATCTCTCCCATCTTACCGAGAAATC 2100
Dh 2041 CATCTCCCTCTGATTTCCCATGCAACTTGCCTTATCTCTCCCATCTTACCGAGAAATC 2100
Qy 2101 TATTTGTTTTTCTTGTGCAATTTTACTATGATGTGTATGTGGCTGTACACCAACCCC 2160
Dh 2101 TATTTGTTTTTCTTGTGCAATTTTACTATGATGTGTATGTGGCTGTACACCAACCCC 2160
Qy 2161 CCCATGGGGGGGTGAGAGGGGTGCAAGGCTGTGCTGTCACTTTTCTACTTTGAA 2220
Dh 2161 CCCATGGGGGGGTGAGAGGGGTGCAAGGCTGTGCTGTCACTTTTCTACTTTGAA 2220
Qy 2221 CTGTATTAATTAATCACTTCTGTTTGTTCAGTTTTCATTAATTAATTAATTAATTAAT 2273
Dh 2221 CTGTATTAATTAATCACTTCTGTTTGTTCAGTTTTCATTAATTAATTAATTAATTAAT 2273

RESULT 2
US-09-880-107-3415
Sequence 3415, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3415
LENGTH: 1842
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 U72515
US-09-880-107-3415

Query Match 81.0%; Score 1842; DB 10; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 TACCTCATCACCCTCTTCCATCCTTTACAGGCTCTCAATGCTATTTTAACTTTGA 338
DB 1 TACCTCATCACCCTCTTCCATCCTTTACAGGCTCTCAATGCTATTTTAACTTTGA 60
QY 339 AACCACTTACACCTCCCTGTGTGTATTTGCTTCACTTCTCACTTCTTGA 398
DB 61 AACCACTTACACCTCCCTGTGTGTATTTGCTTCACTTCTCACTTCTTGA 120
QY 399 GGGCCGACATCACTGCGCTCTCTCACTTCTTCTTCACTTCTTGA 458
DB 121 GGGCCGACATCACTGCGCTCTCTCACTTCTTCTTCACTTCTTGA 180
QY 459 GGATCTATTACACCTGCGCACTGATATGATGATGATGATGATGATGATGATGAT 518
DB 181 GGATCTATTACACCTGCGCACTGATATGATGATGATGATGATGATGATGATGAT 240
QY 519 CTGACTTTGAAGCTGATTTGCTTGTGCTTGAATTTGATGATGATGATGATGATGAT 578
DB 241 CTGACTTTGAAGCTGATTTGCTTGTGCTTGAATTTGATGATGATGATGATGATGAT 300
QY 579 TCTTGTCTCTGAGCAACAGAAATATGCAATGCTGCTTCTTCTTCTGCTGAAAT 638
DB 301 TCTTGTCTCTGAGCAACAGAAATATGCAATGCTGCTTCTTCTTCTGCTGAAAT 360
QY 639 GCTGCTTCTCTCTCTTCTTATGAGGCTCTTGTGATGAGGCTTCTTCTTCTTCTTCT 698
DB 361 GCTGCTTCTCTCTCTTCTTATGAGGCTCTTGTGATGAGGCTTCTTCTTCTTCTTCT 420
QY 699 TACATGAAGCTGTGAGGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 758
DB 421 TACATGAAGCTGTGAGGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 480
QY 759 ATTCTGCTCTCAAGGCTGAGTCTGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 818
DB 481 ATTCTGCTCTCAAGGCTGAGTCTGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
QY 819 ACCCCCAATCAAGAGAGATATCTCTCACTGAGAGATGATGATGATGATGATGATGAT 878
DB 541 ACCCCCAATCAAGAGAGATATCTCTCACTGAGAGATGATGATGATGATGATGATGAT 600
QY 879 TTCGCTGATGTACATGCTGATCTGAGGCAAGTTTGTCTGATGATGATGATGATGAT 938

DB 601 TTCGCTGATGTACATGCTGATCTGAGGCAAGTTTGTCTGATGATGATGATGATGAT 660
QY 938 TGGCTGTACAGAGAGATGATGATTTTGTGAGGCTCTGAGGCTTCAATGCTTTGAAGA 998
DB 661 TGGCTGTACAGAGAGATGATGATTTTGTGAGGCTCTGAGGCTTCAATGCTTTGAAGA 720
QY 999 AAGGCAAGGCAAGTGGGATGCTGCTGCAATGAGAGTGGCTCTTGAACAAAC 1058
DB 721 AAGGCAAGGCAAGTGGGATGCTGCTGCAATGAGAGTGGCTCTTGAACAAAC 780
QY 1059 CCCGCTTCACTGAGCACTTGTCTTCAATCAATCAACCAAGCTGAGTGGCTGAG 1118
DB 781 CCCGCTTCACTGAGCACTTGTCTTCAATCAATCAACCAAGCTGAGTGGCTGAG 840
QY 1119 TACATCTTCAAGCACTCAATGCTTGTGAAATTAAGAACTCTCAAGGCTCTGCTTG 1178
DB 841 TACATCTTCAAGCACTCAATGCTTGTGAAATTAAGAACTCTCTAGGCTCTGCTTG 900
QY 1179 CTATTCCTGAGGCTGAGCAAGGCTGCACTGAGATACCTGCTGCTTCAAGATGAA 1238
DB 901 CTATTCCTGAGGCTTGTGAGCAAGGCTGCACTGAGATACCTGCTGCTTCAAGATGAA 960
QY 1239 TTCTCATTTGATTTGTGAAAGAGAGAGCTGAGCTTCAATCAAGAGAGAGAGAGAG 1298
DB 961 TTCTCATTTGATTTGTGAAAGAGAGAGAGCTGAGCTTCAATCAAGAGAGAGAGAGAG 1020
QY 1299 AGCAAGCTGAGGCACTTGTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1358
DB 1021 AGCAAGCTGAGGCACTTGTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1359 CACTGAGCTTCAAG 1418
DB 1081 CACTGAGCTTCAAG 1140
QY 1419 CTTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1478
DB 1141 CTTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
QY 1479 AATTTGCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1538
DB 1201 AATTTGCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1260
QY 1539 TAAATCATTTCCCTGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
DB 1261 TAAATCATTTCCCTGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1599 ACAGCACTCTTTGAGGAG 1658
DB 1321 ACAGCACTCTTTGAGGAG 1380
QY 1659 TTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1718
DB 1381 TTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1719 TTTGAG 1778
DB 1441 TTTGAG 1500
QY 1779 ACCGCACTCTTCAAGGAG 1838
DB 1501 ACCGCACTCTTCAAGGAG 1560
QY 1839 GGAATTAAGGAG 1898
DB 1561 GGAATTAAGGAG 1620
QY 1899 CACCAAGTGGAG 1958
DB 1621 CACCAAGTGGAG 1680
QY 1959 TCCCAAGTGGAG 2018

Db 1681 TCCCAAGTCTGGATTAAGGCGTGAAGCCACGTCGCCGCCCAAGGGAACACTT 1740
QY 2019 GTGGAGAGACAGAGGGGCTCAATCTCCCTCTGATTCCTCCCATGCACTTCCCTTATC 2078
Db 1741 GTGGAGAGACAGAGGGGCTCAATCTCCCTCTGATTCCTCCCATGCACTTCCCTTATC 1800
QY 2079 TCTCCCATCTAGCAGGAATCTATTGTGTTTCTTCTTCTGSC 2120
Db 1801 TCTCCCATCTAGCAGGAATCTATTGTGTTTCTTCTTCTGSC 1842

RESULT 3
US-09-954-531-1368
Sequence 1368, Application US/09954531
Patent No. US20020165180A1
GENERAL INFORMATION:
APPLICANT: Weaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
TITLE OF INVENTION: Gene Sets
FILE REFERENCE: 689290-77
CURRENT APPLICATION NUMBER: US/09/954,531
CURRENT FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1368
LENGTH: 1842
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-1368

Query Match 81.0%; Score 1842; DB 10; Length 1842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 TACCTCATCACTCTTCCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGA 338
Db 1 TACCTCATCACTCTTCCATACCTTTACAGGCTCTCAATGCTTATTTTAACTTTGA 60
QY 339 AACAGCTCTACACCTCCGCTGTGATTTGCTTCAATCCATCCCTGCACTAAG 398
Db 61 AACAGCTCTACACCTCCGCTGTGATTTGCTTCAATCCATCCCTGCACTAAG 120
QY 399 GGGCGACATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTACCTTGGCT 458
Db 121 GGGCGACATCACTGCGCTCTCACTACCTTTTGTCCAGATGGCTACCTTGGCT 180
QY 459 GGAATCTATTACATGCGACCGGCACTAGATATCAAGTGAACATGCCATTTGTT 518
Db 181 GGAATCTATTACATGCGACCGGCACTAGATATCAAGTGAACATTTGTT 240
QY 519 CTGACTTGAAGCTGATTTGTTGGCTGTGACTTGTGACGGAGGGAAGATCAGAT 578
Db 241 CTGACTTGAAGCTGATTTGTTGGCTGTGACTTGTGACGGAGGGAAGATCAGAT 300
QY 579 TCCCTGTCTCTGAGCAAGAAATATGCAATAGTGTGTTCCCTTCCCTGAGAACT 638
Db 301 TCCCTGTCTCTGAGCAAGAAATATGCAATAGTGTGTTCCCTTCCCTGAGAACT 360
QY 639 GCTGTGTTCTCTACTTCTATGGGCTTTCTGTAGAGGCCCACTTCTCAATGATAC 698
Db 361 GCTGTGTTCTCTACTTCTATGGGCTTTCTGTAGAGGCCCACTTCTCAATGATAC 420
QY 699 TACATGAAGCTGTGAGGAGAGCTGATTTGACATACAGGAAGATACCAACAGCATC 758

Db 421 TACATGAAGCTGTGAGGAGAGAGCTGATTTGACATACAGGAAGATACCAACAGCATC 480
QY 759 ATTCTGCTCTCAAGGCGCTGAGTCTGAGGCTTTTCTTACCTAGTGGCTTACACATGCTC 818
Db 481 ATTCTGCTCTCAAGGCGCTGAGTCTGAGGCTTTTCTTACCTAGTGGCTTACACATGCTC 540
QY 819 AGCCCCCATATCAAGAAATCTATCTCTCACTGAAGACTATAGAACCAACCCCTCTG 878
Db 541 AGCCCCCATATCAAGAAATCTATCTCTCACTGAAGACTATAGAACCAACCCCTCTG 600
QY 879 TTCGCTGATGATCATGTGCTGATCTTGGGCAAGTTTGTCTGATCAAAATATGTCACCTGT 938
Db 601 TTCGCTGATGATCATGTGCTGATCTTGGGCAAGTTTGTCTGATCAAAATATGTCACCTGT 660
QY 939 TGGCTGTGATCAAGAGAGATGATTTTGAAGGCTGAGGCTTCAATGCTTTGAAGAA 998
Db 661 TGGCTGTGATCAAGAGAGATGATTTTGAAGGCTGAGGCTTCAATGCTTTGAAGAA 720
QY 999 AAGGCAAGCAAGTGGGATGCTGTGCAACATGAAGTGTGCTTTTGAACAAAC 1058
Db 721 AAGGCAAGCAAGTGGGATGCTGTGCAACATGAAGTGTGCTTTTGAACAAAC 780
QY 1059 CCGGCTTCACTGACACCAATGCTCATTCACATCAACCAACGCTGGTGGCCGC 1118
Db 781 CCGGCTTCACTGACACCAATGCTCATTCACATCAACCAACGCTGGTGGCCGC 840
QY 1119 TACATCTTCAAGACTCAAGTCTTGGAAATTAAGAACTCTCTGAGGCTCTGCTG 1178
Db 841 TACATCTTCAAGACTCAAGTCTTGGAAATTAAGAACTCTCTGAGGCTCTGCTG 900
QY 1179 CTATTCCTGCGCTCTGGAAGGCGCTGCACTAGATACCTGCTCTTCCAGATGAA 1238
Db 901 CTATTCCTGCGCTCTGGAAGGCGCTGCACTAGATACCTGCTCTTCCAGATGAA 960
QY 1239 TTCTCATTTGATTTGTGAAAGACAGGCTGCAAGCTCATTCAGAGAGCCCACTG 1298
Db 961 TTCTCATTTGATTTGTGAAAGACAGGCTGCAAGCTCATTCAGAGAGCCCACTG 1020
QY 1299 AGCAAGCTGCGCGCAATTAAGTCTCTCAAGCTTCAATTTGTTGTCACAGACATC 1358
Db 1021 AGCAAGCTGCGCGCAATTAAGTCTCTCAAGCTTCAATTTGTTGTCACAGACATC 1080
QY 1359 CACTGCTCTTCAAGGCTTCACTGCAAGTCTTCTGCTCTTCAAGTGGCAATAG 1418
Db 1081 CACTGCTCTTCAAGGCTTCACTGCAAGTCTTCTGCTCTTCAAGTGGCAATAG 1140
QY 1419 CTAAAGTGTAAATCAATCTATTTCTTGGCACAATCTTCTCTGAGCTACTATTC 1478
Db 1141 CTAAAGTGTAAATCAATCTATTTCTTGGCACAATCTTCTCTGAGCTACTATTC 1200
QY 1479 ATATGCTTATTAATCAAGAGATGTCAGAGAAAGAAAGTAAAGAAATGAA 1538
Db 1201 ATATGCTTATTAATCAAGAGATGTCAGAGAAAGAAAGTAAAGAAATGAA 1260
QY 1539 TATTCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1598
Db 1261 TATTCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1599 ACAGCACTCTTGGCCCGAGAGAGAAATGAAAGCAAGGAGGTGGAAGTCAATG 1658
Db 1321 ACAGCACTCTTGGCCCGAGAGAGAAATGAAAGCAAGGAGGTGGAAGTCAATG 1380
QY 1659 TTCCAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1718
Db 1381 TTCCAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1719 TTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1778
Db 1441 TTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1779 ACCGAACTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1838

Db 1501 ACCGCAACCTCCACCTCTGGGTTCAAGTATTTTCCTGCTCAGCTCCACGATAGCTG 1560
QY 1839 GGAATACAGGACGCGCCACCATGCGCAGTAATTTTGTATTTTACGTAGAAAGGGATTT 1898
Db 1561 GGAATACAGGACGCGCCACCATGCGCAGTAATTTTGTATTTTACGTAGAAAGGGATTT 1620
QY 1899 CACCAAGTGGCCAGGCTGGTCTCGAACTCTGACCGCAAGTATCCACCCCTCCGCG 1958
Db 1621 CACCAAGTGGCCAGGCTGGTCTCGAACTCTGACCGCAAGTATCCACCCCTCCGCG 1680
QY 1959 TCCCAAGTGGCCAGGCTGGTCTCGAACTCTGACCGCAAGTATCCACCCCTCCGCG 2018
Db 1681 TCCCAAGTGGCCAGGCTGGTCTCGAACTCTGACCGCAAGTATCCACCCCTCCGCG 1740
QY 2019 GTGGAGGAGCAGAGGGGCTCATCTCCCTCTGATTTCCCGCAGCAATTGCTTATC 2078
Db 1741 GTGGAGGAGCAGAGGGGCTCATCTCCCTCTGATTTCCCGCAGCAATTGCTTATC 1800
QY 2079 TCTCCCATCTACGACGAGATCTATTTGTTTCTTCTGCC 2120
Db 1801 TCTCCCATCTACGACGAGATCTATTTGTTTCTTCTGCC 1842

RESULT 4
US-09-764-860-1135/c
; Sequence 1135, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1135
; LENGTH: 7461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1135

Query Match 31.1%; Score 708; DB 9; Length 7461;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GGTGGCTGTGTGGGAGCTGTGTCAGAACTACTGCTCTCTTTCACAGCACTCCTTTG 1612
Db 714 GGTGGCTGTGTGGGAGCTGTGTCAGAACTACTGCTCTCTTTCACAGCACTCCTTTG 655
QY 1613 CCCGAGCAGAGAGTGAAGAAAGCCAGGAGGTGGAAGATGATGCTTCCAGCTGTGCT 1672
Db 654 CCCGAGCAGAGAGTGAAGAAAGCCAGGAGGTGGAAGATGATGCTTCCAGCTGTGCT 595
QY 1673 CTGCTGCCAGCAAGCTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGAGAAAGCGT 1732
Db 594 CTGCTGCCAGCAAGCTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGAGAAAGCGT 535
QY 1733 CTGCTTTGTCAACCAAGCTGGAATGCAATGCGGGAGTTCAGCTCACCGCACTCCAC 1792
Db 534 CTGCTTTGTCAACCAAGCTGGAATGCAATGCGGGAGTTCAGCTCACCGCACTCCAC 475
QY 1793 CTCCTGGGTTCAAGTATTTTCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG 1852
Db 474 CTCCTGGGTTCAAGTATTTTCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG 415
QY 1853 CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGATTTCAACAGTTGGCCA 1912
Db 414 CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGATTTCAACAGTTGGCCA 355
QY 1913 GGCTGGTCTCGAACTCTGACCGCAAGTATCCACCGCTCCGCTCCCAAGTGTCTGG 1972
Db 354 GGCTGGTCTCGAACTCTGACCGCAAGTATCCACCGCTCCGCTCCCAAGTGTCTGG 295

QY 1973 GATTACAGCGGTGAGCCACCGTGCCTCCGCCCCAAAGGGGAACTCTGTGGAGAGCAGA 2032
Db 294 GATTACAGCGGTGAGCCACCGTGCCTCCGCCCCAAAGGGGAACTCTGTGGAGAGCAGA 235
QY 2033 GGGGCTCAATCTCCCTCTGATTTCCCGCAGCAATTGCTTATCTCTCCCATCTAGC 2092
Db 234 GGGGCTCAATCTCCCTCTGATTTCCCGCAGCAATTGCTTATCTCTCCCATCTAGC 175
QY 2093 CAGGAATCTATTTGTTTCTTCTGCGCAATTACTATGATTTGTATGTGCGCTACCA 2152
Db 174 CAGGAATCTATTTGTTTCTTCTGCGCAATTACTATGATTTGTATGTGCGCTACCA 115
QY 2153 CCACCCCGCCAGTGGGGGGGTGAGAGGGGGTGCAGAGGCGCTGCGCTCACTTTTCTA 2212
Db 114 CCACCCCGCCAGTGGGGGGGTGAGAGGGGGTGCAGAGGCGCTGCGCTCACTTTTCTA 55
QY 2213 CTTTGAACGTATATAGTAAATCATCTTCTGTTTGTTCAGTTTTC 2260
Db 54 CTTTGAACGTATATAGTAAATCATCTTCTGTTTGTTCAGTTTTC 7

RESULT 5
US-09-764-904-91/c
; Sequence 91, Application US/09764904
; Patent No. US20020173454A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764,904
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 7461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-904-91

Query Match 31.1%; Score 708; DB 10; Length 7461;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GGTGGCTGTGTGGGAGCTGTGTCAGAACTACTGCTCTCTTTCACAGCACTCCTTTG 1612
Db 714 GGTGGCTGTGTGGGAGCTGTGTCAGAACTACTGCTCTCTTTCACAGCACTCCTTTG 655
QY 1613 CCCGAGCAGAGAGTGAAGAAAGCCAGGAGGTGGAAGATGATGCTTCCAGCTGTGCT 1672
Db 654 CCCGAGCAGAGAGTGAAGAAAGCCAGGAGGTGGAAGATGATGCTTCCAGCTGTGCT 595
QY 1673 CTGCTGCCAGCAAGCTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGAGAAAGCGT 1732
Db 594 CTGCTGCCAGCAAGCTCTTCAATTTGGGGCCAAAGGGGAACTTTTGTGAGAAAGCGT 535
QY 1733 CTGCTTTGTCAACCAAGCTGGAATGCAATGCGGGAGTTCAGCTCACCGCACTCCAC 1792
Db 534 CTGCTTTGTCAACCAAGCTGGAATGCAATGCGGGAGTTCAGCTCACCGCACTCCAC 475
QY 1793 CTCCTGGGTTCAAGTATTTTCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG 1852
Db 474 CTCCTGGGTTCAAGTATTTTCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG 415
QY 1853 CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGATTTCAACAGTTGGCCA 1912
Db 414 CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGATTTCAACAGTTGGCCA 355
QY 1913 GGCTGGTCTCGAACTCTGACCGCAAGTATCCACCGCTCCGCTCCCAAGTGTCTGG 1972
Db 354 GGCTGGTCTCGAACTCTGACCGCAAGTATCCACCGCTCCGCTCCCAAGTGTCTGG 295

```
QY 1973 GATTACAGGCGTGAAGCCACCGTGCCTCCGCCCCAAAGGGGAAACTCTTGTGGAGAGACAGA 2032
      |||
Db 294 GATTACAGGCGTGAAGCCACCGTGCCTCCGCCCCAAAGGGGAAACTCTTGTGGAGAGACAGA 235
QY 2033 GGGGCTCACATCTCCCTCTGATTTCCCCATGACATTCCTTATCTCTCCCATCTAGC 2092
      |||
Db 234 GGGGCTCACATCTCCCTCTGATTTCCCCATGACATTCCTTATCTCTCCCATCTAGC 175
QY 2093 CAGGAATCTAATGTGTCTTCTTCTGCAATTTACTATGATGTGTATGTGCGCTACCA 2152
      |||
Db 174 CAGGAATCTAATGTGTCTTCTTCTGCAATTTACTATGATGTGTATGTGCGCTACCA 115
QY 2153 CCACCCCCCATGAGGGGGGTGAGAGGGGTGCAAGGCGCTGCTGCCTCACTTTTCTA 2212
      |||
Db 114 CCACCCCCCATGAGGGGGGTGAGAGGGGTGCAAGGCGCTGCTGCCTCACTTTTCTA 55
QY 2213 CCTTGAAGCTGTATGATTAATAATCACTTCTGTTTGTTCAGTTTCTCA 2260
      |||
Db 54 CCTTGAAGCTGTATGATTAATAATCACTTCTGTTTGTTCAGTTTCTCA 7
```

```
RESULT 6
US-10-091-548-91/c
; Sequence 91, Application US/10091548
; Publication No. US20030049703A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122C1
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ. ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ. ID NO 91
; LENGTH: 7461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-91
```

```
Query Match 31.1%; Score 708; DB 14; Length 7461;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GGTGGCTGTGGCGGAGCTGTGCGAGAACTACTGTCTCCCTTTTCACAGACTCTTTG 1612
      |||
Db 714 GGTGGCTGTGGCGGAGCTGTGCGAGAACTACTGTCTCCCTTTTCACAGACTCTTTG 655
QY 1613 CCCGAGCAGAGAAATGAAAAGCCAGGAGGTGGAAGATGATCTTCCAGCTGTGCT 1672
      |||
Db 654 CCCGAGCAGAGAAATGAAAAGCCAGGAGGTGGAAGATGATCTTCCAGCTGTGCT 595
QY 1673 CTGCTGCGAGCAAGTCTTCAATTTGGGCGCAAGGGGAACTTTTTTTGGAGAAGCGT 1732
      |||
Db 594 CTGCTGCGAGCAAGTCTTCAATTTGGGCGCAAGGGGAACTTTTTTTGGAGAAGCGT 535
QY 1733 CTGCTGTTGTCAACCGAGCTGGAATGCAATGCGGGATCTCAAGTCAACCGCAACTCCAC 1792
      |||
Db 534 CTGCTGTTGTCAACCGAGCTGGAATGCAATGCGGGATCTCAAGTCAACCGCAACTCCAC 475
QY 1793 CTCTGGGTTCAAGTATTTTCCGCTCAGGCTCCCAAGTATGGAATATACAGGACG 1852
      |||
Db 474 CTCTGGGTTCAAGTATTTTCCGCTCAGGCTCCCAAGTATGGAATATACAGGACG 415
QY 1853 CCACCATGCGCAGCTAATTTTGTATTTTCATAGAAAAGGATTTTCAACAGTTGGCA 1912
      |||
Db 414 CCACCATGCGCAGCTAATTTTGTATTTTCATAGAAAAGGATTTTCAACAGTTGGCA 355
QY 1913 GGTGTGTCTGAATCTCTGACCGCAAGTATCCACCGGCTCTCCGCTCCCAAGTGTGG 1972
      |||
Db 354 GGTGTGTCTGAATCTCTGACCGCAAGTATCCACCGGCTCTCCGCTCCCAAGTGTGG 295
QY 1973 GATTACAGGCGTGAAGCCACCGTGCCTCCGCCCCAAAGGGGAAACTCTTGTGGAGAGACAGA 2032
```

```
Db 294 GATTACAGGCGTGAAGCCACCGTGCCTCCGCCCCAAAGGGGAAACTCTTGTGGAGAGACAGA 235
QY 2033 GGGGCTCACATCTCCCTCTGATTTCCCCATGACATTCCTTATCTCTCCCATCTAGC 2092
      |||
Db 234 GGGGCTCACATCTCCCTCTGATTTCCCCATGACATTCCTTATCTCTCCCATCTAGC 175
QY 2093 CAGGAATCTAATGTGTCTTCTTCTGCAATTTACTATGATGTGTATGTGCGCTACCA 2152
      |||
Db 174 CAGGAATCTAATGTGTCTTCTTCTGCAATTTACTATGATGTGTATGTGCGCTACCA 115
QY 2153 CCACCCCCCATGAGGGGGGTGAGAGGGGTGCAAGGCGCTGCTGCCTCACTTTTCTA 2212
      |||
Db 114 CCACCCCCCATGAGGGGGGTGAGAGGGGTGCAAGGCGCTGCTGCCTCACTTTTCTA 55
QY 2213 CCTTGAAGCTGTATGATTAATAATCACTTCTGTTTGTTCAGTTTCTCA 2260
      |||
Db 54 CCTTGAAGCTGTATGATTAATAATCACTTCTGTTTGTTCAGTTTCTCA 7
```

```
RESULT 7
US-10-074-095-1135/c
; Sequence 1135, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
```

```

PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22

```

```

PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475

```

```
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match      31.1%; Score 708; DB 14; Length 7461;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GGTGGCCCTGTGGCGGAGTGTGCGAGAACTACTGCTGCTCTCTTTTCACAGCACTCTTTG 1612
DB 714 GGTGGCCCTGTGGCGGAGTGTGCGAGAACTACTGCTGCTCTCTTTTCACAGCACTCTTTG 655
QY 1613 CCCAGAGCAGAGATGGAAGAACCGAGGAGGAGGAGATGATGCTTCAGCTGAGCGCT 1672
DB 654 CCCAGAGCAGAGATGGAAGAACCGAGGAGGAGGAGATGATGCTTCAGCTGAGCGCT 595
QY 1673 CTGCTGCGACCAAGTCTTCACTTTGGGGCAAGGGGAACTTTTGGAGAGGCGT 1732
DB 594 CTGCTGCGACCAAGTCTTCACTTTGGGGCAAGGGGAACTTTTGGAGAGGCGT 535
QY 1733 CTGCTGCTTGTACCCAGCTGGAATGAGTGGGGGATCTCAGCTCAGCGCACTTCAC 1792
DB 534 CTGCTGCTTGTACCCAGCTGGAATGAGTGGGGGATCTCAGCTCAGCGCACTTCAC 475
QY 1793 CTGCTGCGTCAAGTATTTCTGCTCCTCAGCTCCCAAGTATGCTGGGATATAGAGCAG 1852
DB 474 CTGCTGCGTCAAGTATTTCTGCTCCTCAGCTCCCAAGTATGCTGGGATATAGAGCAG 415
QY 1853 CCACATGCCCCAGCTAATTTTGTATTTTCAAGTAAAGCGGATTTTACACCGTTGGCCA 1912
DB 414 CCACATGCCCCAGCTAATTTTGTATTTTCAAGTAAAGCGGATTTTACACCGTTGGCCA 355
QY 1913 GCGTGTCTGGAATCTCTGACCCGCAAGTATCCACCCGCTCCGCTCCCAAGTGTGG 1972
DB 354 GCGTGTCTGGAATCTCTGACCCGCAAGTATCCACCCGCTCCGCTCCCAAGTGTGG 295
QY 1973 GATTACAGGGGTGAGCCAGCGTGCCTGGCCCAAGGGGAACTTGTGGAGGAGAGAGA 2032
DB 294 GATTACAGGGGTGAGCCAGCGTGCCTGGCCCAAGGGGAACTTGTGTGGAGGAGAGAGA 235
QY 2033 GGGGCTCAGATCTCCCTCTGATTTCCCCATGACATGTGCTTATCTCTCCCATCTAGC 2092
DB 234 GGGGCTCAGATCTCCCTCTGATTTCCCCATGACATGTGCTTATCTCTCCCATCTAGC 175
QY 2093 CAGGATCTATGTGTTTTTCTTCTGCAATTTATCTATGATGTAATGTCGCGCTACCA 2152
DB 174 CAGGATCTATGTGTTTTTCTTCTGCAATTTATCTATGATGTAATGTCGCGCTACCA 115
QY 2153 CCAACCCCCCAGTGGGGGGGAGAGAGGGGTGCAAGGCTGCTGCTGCTCACTTTTCTA 2212
DB 114 CCAACCCCCCAGTGGGGGGGAGAGAGGGGTGCAAGGCTGCTGCTGCTCACTTTTCTA 55
QY 2213 CTTTGAAGTGTATTAATAAATCACTTCTGTTTGTTCAGTTTTC 2260
DB 54 CTTTGAAGTGTATTAATAAATCACTTCTGTTTGTTCAGTTTTC 7

RESULT 8
US-09-908-711-158/c
; Sequence 158, Application US/09908711
; Patent No. US2002045230A1
; GENERAL INFORMATION:
; APPLICANT: Roese et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA128
; CURRENT APPLICATION NUMBER: US/09/908,711
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US01/01360
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,867
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01344
; PRIOR FILING DATE: 2001-01-17
```

```
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01345
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,905
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01354
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01339
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01340
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,874
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01334
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,898
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01320
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,853
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01349
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,902
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01239
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01348
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,882
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01347
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,896
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01307
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01341
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,856
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01336
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,868
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01312
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 158
; LENGTH: 32174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

NAME/KEY: SITE	LOCATION: (29356)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29357)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29358)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29359)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29360)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29361)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29362)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29363)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29364)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29365)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29366)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29367)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29368)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29369)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29370)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29371)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29372)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29373)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29374)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29375)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29376)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29377)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29378)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29379)	OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE	LOCATION: (29380)	OTHER INFORMATION: n equals a,t,g, or c

```

1 LOCATION: (29380)
2 OTHER INFORMATION: n equals a,t,g, or c
3 NAME/KEY: SITE
4 LOCATION: (29381)
5 OTHER INFORMATION: n equals a,t,g, or c
6 NAME/KEY: SITE
7 LOCATION: (29382)
8 OTHER INFORMATION: n equals a,t,g, or c
9 NAME/KEY: SITE
10 LOCATION: (29383)
11 OTHER INFORMATION: n equals a,t,g, or c
12 NAME/KEY: SITE
13 LOCATION: (29384)
14 OTHER INFORMATION: n equals a,t,g, or c
15 NAME/KEY: SITE
16 LOCATION: (29385)
17 OTHER INFORMATION: n equals a,t,g, or c
18 NAME/KEY: SITE
19 LOCATION: (29386)
20 OTHER INFORMATION: n equals a,t,g, or c
21 NAME/KEY: SITE
22 LOCATION: (29387)
23 OTHER INFORMATION: n equals a,t,g, or c
24 NAME/KEY: SITE
25 LOCATION: (29388)
26 OTHER INFORMATION: n equals a,t,g, or c
27 NAME/KEY: SITE
28 LOCATION: (29389)
29 OTHER INFORMATION: n equals a,t,g, or c
30 NAME/KEY: SITE
31 LOCATION: (29390)
32 OTHER INFORMATION: n equals a,t,g, or c
33 NAME/KEY: SITE
34 LOCATION: (29391)
35 OTHER INFORMATION: n equals a,t,g, or c
36 NAME/KEY: SITE
37 LOCATION: (29392)
38 OTHER INFORMATION: n equals a,t,g, or c
39 NAME/KEY: SITE
40 LOCATION: (29393)

```

Query Match 31.1%; Score 708; DB 9; Length 32174;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 708; Conservative 0; Mismatches 0; Gaps 0;

QY	1553	GGTGGCCTGGCGGGACGTGGGAGCAAACTACTGCTCCCTTTTCAACAGACTCCTTGG	16112
Db	6043	GGTGGCCTGGCGGGACGTGGGAGCAAACTACTGCTCCCTTTTCAACAGACTCCTTGG	59894
QY	1613	CCCCAGACGACGATATGAAAAGCCAGGAGAGTGGAAATGATATGATCTTCAGCTGTGCTCT	16722
Db	5983	CCCCAGACGACGATATGAAAAGCCAGGAGAGTGGAAATGATATGATATGATCTTCAGCTGTGCTCT	5924
QY	1673	CTGCTGCACGCAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTTTTGGAGAGGCGT	17322
Db	5923	CTGCTGCACGCAAGTCTTCAATTTGGGGCCAAAGGGGAAACTTTTTTTGGAGAGGCGT	5864
QY	1733	CTTGGCTTTGTACCCACGCTGGAATGAGTGGGGGAACTCAGCTCACCCCAACTCTCAC	1792
Db	5863	CTTGGCTTTGTACCCACGCTGGAATGAGTGGGGGAACTCAGCTCACCCCAACTCTCAC	5804
QY	1793	CTCCTGGGTTCAAGTATTTTTCCTGCTCAGCCTCCCAAGTAGCTGGGAAATACAGGACG	1852
Db	5803	CTCCTGGGTTCAAGTATTTTTCCTGCTCAGCCTCCCAAGTAGCTGGGAAATACAGGACG	5744
QY	1853	CCACCATGCCAGCTAATTTTGTATTTTCAGTATGAAAAGGGAGTTTCAACAAGTGGCCA	1912
Db	5743	CCACCATGCCAGCTAATTTTGTATTTTCAGTATGAAAAGGGAGTTTCAACAAGTGGCCA	5684
QY	1913	GGCTGGCTGGAACCTCGTACCGCAATGATATCCACCGGCTCCGGCTCCCAAAAGTGTGG	1972
Db	5683	GGCTGGCTGGAACCTCGTACCGCAATGATATCCACCGGCTCCGGCTCCCAAAAGTGTGG	5624

1	OTHER INFORMATION:	n	equals	a,t,g,	or c
2	NAME/KEY: SITE				
3	LOCATION: (29392)				
4	OTHER INFORMATION:	n	equals	a,t,g,	or c
5	NAME/KEY: SITE				
6	LOCATION: (29393)				
7	OTHER INFORMATION:	n	equals	a,t,g,	or c
8	NAME/KEY: SITE				
9	LOCATION: (29394)				
10	OTHER INFORMATION:	n	equals	a,t,g,	or c
11	NAME/KEY: SITE				
12	LOCATION: (29395)				
13	OTHER INFORMATION:	n	equals	a,t,g,	or c
14	NAME/KEY: SITE				
15	LOCATION: (29396)				
16	OTHER INFORMATION:	n	equals	a,t,g,	or c
17	NAME/KEY: SITE				
18	LOCATION: (29397)				
19	OTHER INFORMATION:	n	equals	a,t,g,	or c
20	NAME/KEY: SITE				
21	LOCATION: (29398)				
22	OTHER INFORMATION:	n	equals	a,t,g,	or c
23	NAME/KEY: SITE				
24	LOCATION: (29399)				
25	OTHER INFORMATION:	n	equals	a,t,g,	or c
26	NAME/KEY: SITE				
27	LOCATION: (29400)				
28	OTHER INFORMATION:	n	equals	a,t,g,	or c
29	NAME/KEY: SITE				
30	LOCATION: (29401)				
31	OTHER INFORMATION:	n	equals	a,t,g,	or c
32	NAME/KEY: SITE				
33	LOCATION: (29402)				
34	OTHER INFORMATION:	n	equals	a,t,g,	or c
35	NAME/KEY: SITE				
36	LOCATION: (29403)				
37	OTHER INFORMATION:	n	equals	a,t,g,	or c
38	NAME/KEY: SITE				
39	LOCATION: (29404)				
40	OTHER INFORMATION:	n	equals	a,t,g,	or c
41	NAME/KEY: SITE				
42	LOCATION: (29405)				
43	OTHER INFORMATION:	n	equals	a,t,g,	or c
44	NAME/KEY: SITE				
45	LOCATION: (29406)				
46	OTHER INFORMATION:	n	equals	a,t,g,	or c
47	NAME/KEY: SITE				
48	LOCATION: (29407)				
49	OTHER INFORMATION:	n	equals	a,t,g,	or c
50	NAME/KEY: SITE				
51	LOCATION: (29408)				
52	OTHER INFORMATION:	n	equals	a,t,g,	or c
53	NAME/KEY: SITE				
54	LOCATION: (29409)				
55	OTHER INFORMATION:	n	equals	a,t,g,	or c
56	NAME/KEY: SITE				
57	LOCATION: (29410)				
58	OTHER INFORMATION:	n	equals	a,t,g,	or c
59	NAME/KEY: SITE				
60	LOCATION: (29411)				
61	OTHER INFORMATION:	n	equals	a,t,g,	or c
62	NAME/KEY: SITE				
63	LOCATION: (29412)				
64	OTHER INFORMATION:	n	equals	a,t,g,	or c
65	NAME/KEY: SITE				
66	LOCATION: (29413)				
67	OTHER INFORMATION:	n	equals	a,t,g,	or c
68	NAME/KEY: SITE				
69	LOCATION: (29414)				
70	OTHER INFORMATION:	n	equals	a,t,g,	or c
71	NAME/KEY: SITE				
72	LOCATION: (29415)				
73	OTHER INFORMATION:	n	equals	a,t,g,	or c

```

; NAME/KEY: SITE
; LOCATION: (29416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

Query Match      31.1%; Score 708; DB 9; Length 32174;
Best Local Similarity 100.0%; Pzed. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      1553  GGTGGCCCTGTGCGGGAGCTGTGTGCAGAACTACTGCTCTCCCTTTTTCACAGCACTCTTTG 1612
        GGTGGCCCTGTGCGGGAGCTGTGTGCAGAACTACTGCTCTCTCCCTTTTTCACAGCACTCTTTG 5984
        6043  GGTGGCCCTGTGCGGGAGCTGTGTGCAGAACTACTGCTCTCTCCCTTTTTCACAGCACTCTTTG 5984

Db      1613  CCCCAGACAGAGAATGAAAAAGCCAGGAGGTGGAAGATGATGATGCTTCCAGCTGTGCT 1672
        CCCCAGACAGAGAATGAAAAAGCCAGGAGGTGGAAGATGATGATGCTTCCAGCTGTGCT 5924
        5983  CCCCAGACAGAGAATGAAAAAGCCAGGAGGTGGAAGATGATGATGCTTCCAGCTGTGCT 5924

Db      1673  CTGCTGCCAGCCCAAGTCTTCAATTTTGGGGCCCAAGGGGAACTTTTTTTTGGAGAAAGCGT 1732
        CTGCTGCCAGCCCAAGTCTTCAATTTTGGGGCCCAAGGGGAACTTTTTTTTGGAGAAAGCGT 5864
        5923  CTGCTGCCAGCCCAAGTCTTCAATTTTGGGGCCCAAGGGGAACTTTTTTTTGGAGAAAGCGT 5864

Db      1733  CTGCTTTTGTACCCAGCAGTGTGAATGAGTGGCGGATCTCAGCTCAACCGCACTCTCAC 1792
        CTGCTTTTGTACCCAGCAGTGTGAATGAGTGGCGGATCTCAGCTCAACCGCACTCTCAC 5804
        5863  CTGCTTTTGTACCCAGCAGTGTGAATGAGTGGCGGATCTCAGCTCAACCGCACTCTCAC 5804

Db      1793  CTCCTGGGTTCAAGTATTTTCTGCTCAGCCCTCCCAAGTGTGGGAATACAGGCAG 1852
        CTCCTGGGTTCAAGTATTTTCTGCTCAGCCCTCCCAAGTGTGGGAATACAGGCAG 5744
        5803  CTCCTGGGTTCAAGTATTTTCTGCTCAGCCCTCCCAAGTGTGGGAATACAGGCAG 5744

Db      1853  CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAACGGGATTTTCAACCACTTGCGCA 1912
        CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAACGGGATTTTCAACCACTTGCGCA 5684
        5743  CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAACGGGATTTTCAACCACTTGCGCA 5684

Db      1913  GGTGTGTCTGGAATCTCCGACCCGAGAGTATCCACCCGCTCCGCTCCCAAGTGTG 1972
        GGTGTGTCTGGAATCTCCGACCCGAGAGTATCCACCCGCTCCGCTCCCAAGTGTG 5624
        5683  GGTGTGTCTGGAATCTCCGACCCGAGAGTATCCACCCGCTCCGCTCCCAAGTGTG 5624

Db      1973  GATTACAGGCGTGAACCAACGCTGCCCCGACCAAGGGGAACTTTGTGGAGAGACAGA 2032
        GATTACAGGCGTGAACCAACGCTGCCCCGACCAAGGGGAACTTTGTGGAGAGACAGA 5564
        5623  GATTACAGGCGTGAACCAACGCTGCCCCGACCAAGGGGAACTTTGTGGAGAGACAGA 5564

Db      2033  GGGGCTCACATCTCCCTCTGTATTTCCCCATGACATTTGCTTATCTTCCCATCTAGC 2092
        GGGGCTCACATCTCCCTCTGTATTTCCCCATGACATTTGCTTATCTTCCCATCTAGC 5504
        5563  GGGGCTCACATCTCCCTCTGTATTTCCCCATGACATTTGCTTATCTTCCCATCTAGC 5504

Db      2093  CAGGAATCTATGTGTTTTCTTCTGCAATTTACTATGATGTGTAATGTCGCGCTACCA 2152
        CAGGAATCTATGTGTTTTCTTCTGCAATTTACTATGATGTGTAATGTCGCGCTACCA 5444
        5503  CAGGAATCTATGTGTTTTCTTCTGCAATTTACTATGATGTGTAATGTCGCGCTACCA 5444

Db      2153  CCACCCCCCAGTGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTGTCCACTTTTCTA 2212
        CCACCCCCCAGTGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTGTCCACTTTTCTA 5384
        5443  CCACCCCCCAGTGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTGTCCACTTTTCTA 5384

Db      2213  CCTTGAACGTGTATTAGATAAATCACTCTCTGTTTGTTCAGTTTTTCA 2260
        CCTTGAACGTGTATTAGATAAATCACTCTCTGTTTGTTCAGTTTTTCA 5336
        5383  CCTTGAACGTGTATTAGATAAATCACTCTCTGTTTGTTCAGTTTTTCA 5336

RESULT 10
US-09-764-877-2645/c
; Sequence 2645, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2645
; LENGTH: 32174

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (29356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29357)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29358)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29360)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29361)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29362)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29363)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29364)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29365)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29366)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29367)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29368)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29369)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29370)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29371)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29372)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29373)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29374)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29375)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29376)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29377)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29378)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29403)
OTHER INFORMATION: n equals a,t,g, or c

```

```
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29404)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29405)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29406)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29407)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29408)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29409)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29410)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29411)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29412)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29413)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29414)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29415)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29416)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
```

```
Query Match 31.1%; Score 708; DB 10; Length 32174;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1553 GGTGGCCCTGGGGGAGCTGTGTCAGAACTACTCCCTTTTACAGCACTCTTTG 1612
DB 6043 GGTGGCCCTGGGGGAGCTGTGTCAGAACTACTCCCTTTTACAGCACTCTTTG 5984
QY 1613 CCCAGAGCAGAGAAATGAAAAAGCCAGGAGGTGAAAGATCGATGCTTCCAGCTGGCCT 1672
DB 5983 CCCAGAGCAGAGAAATGAAAAAGCCAGGAGGTGAAAGATCGATGCTTCCAGCTGGCCT 5924
QY 1673 CTGCTGCCAGCAAGCTTTTCATTTGGGGCCAAAGGGGAACTTTTTTTTGAGAAAGGCGT 1732
DB 5923 CTGCTGCCAGCAAGCTTTTCATTTGGGGCCAAAGGGGAACTTTTTTTTGAGAAAGGCGT 5864
QY 1733 CTTCCTTTTGACCCACGCTGGAATGCAATGCGGAGATCTCAGCTCACCGCAACTCCAC 1792
DB 5863 CTTCCTTTTGACCCACGCTGGAATGCAATGCGGAGATCTCAGCTCACCGCAACTCCAC 5804
QY 1793 CTTCCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGTGGGGAATACAGGCAAG 1852
DB 5803 CTTCCTGGGTTCAAGTATTTTCTGCTCAGCCTCCCAAGTGGGGAATACAGGCAAG 5744
QY 1853 CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGAATTTTCAACAGTTGGCCA 1912
DB 5743 CCACCATGCCAGCTAATTTTGTATTTTCAGTAGAAAGGGAATTTTCAACAGTTGGCCA 5684
QY 1913 GGCTGTCTTCGAATCTCTTACCCGCAAGTATCCACCCGCTCCGCTCCCAAGTGGCTG 1972
DB 5683 GGCTGTCTTCGAATCTCTTACCCGCAAGTATCCACCCGCTCCGCTCCCAAGTGGCTG 5624
```

```
QY 1973 GATTACAGCGTGAAGCCACCGGCGCCGCAAAAGGGAATCTTTGGAGAGAGCAGA 2032
DB 5623 GATTACAGCGTGAAGCCACCGGCGCCGCAAAAGGGAATCTTTGGAGAGAGCAGA 5564
QY 2033 GGGGCTCACATCTCCCTCTGATTCGCCCATGACATTGCTCTCCCATCTAGC 2092
DB 5563 GGGGCTCACATCTCCCTCTGATTCGCCCATGACATTGCTCTCCCATCTAGC 5504
QY 2093 CAGGATCTATTGTTGTTTCTTCTGCGCAATTACTATGATGATGATGCGGCTACCA 2152
DB 5503 CAGGATCTATTGTTGTTTCTTCTGCGCAATTACTATGATGATGATGCGGCTACCA 5444
QY 2153 CCACCCCGCCCATGGGGGGGTGAGAGAGGGGTSCAAGGCCCTGCTCCACTTTTCTA 2212
DB 5443 CCACCCCGCCCATGGGGGGGTGAGAGAGGGGTSCAAGGCCCTGCTCCACTTTTCTA 5384
QY 2213 CTTTGAACGTGATTAGATAAATGACCTTCTGTTGTTCACTTTTCA 2260
DB 5383 CTTTGAACGTGATTAGATAAATGACCTTCTGTTGTTCACTTTTCA 5316

RESULT 11
US-09-670-232/c
Sequence 232, Application US/09860670
Patent No. US20020165137A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA127P1
CURRENT APPLICATION NUMBER: US/09/860,670
PRIORITY FILING DATE: 2001-05-21
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 289
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 232
LENGTH: 32174
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (29356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29357)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29358)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29360)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29361)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29362)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29363)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29364)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29365)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29366)
OTHER INFORMATION: n equals a,t,g, or c
```

LOCATION: (29391)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29392)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29393)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29394)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29397)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29398)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29399)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29400)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29401)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29402)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29403)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29404)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29405)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29406)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29407)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29408)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29409)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29410)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29411)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29412)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29413)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29414)	n	equals	a,t,g,	or	c
OTHER INFORMATION:					
NAME/KEY: SITE					
LOCATION: (29415)	n	equals	a,t,g,	or	c

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29416)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE

Query Match 31.1%; Score 708; DB 10; Length 32174;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GGTGGCTGTGGCGGAGCTGTGCAAGAACTGCTGCTCTCTTTTTCACAGCACTCTCTTGG 1612
DB 6043 GGTGGCTGTGGCGGAGCTGTGCAAGAACTGCTGCTCTCTTTTTCACAGCACTCTCTTGG 5984
QY 1613 CCCAG 1672
DB 5983 CCCAG 5924
QY 1673 CTGGCTCCAGCCAGCTCTTCATTTGGGGCCAAAGGGGAACTTTTGGAGAGAGCGT 1732
DB 5923 CTGGCTCCAGCCAGCTCTTCATTTGGGGCCAAAGGGGAACTTTTGGAGAGAGCGT 5864
QY 1733 CTGGCTTTGTCAACCAAGCTGTGAGTCAAGTGCAGATCTGACCTCACCGCAACTCTCAC 1792
DB 5863 CTGGCTTTGTCAACCAAGCTGTGAGTCAAGTGCAGATCTGACCTCACCGCAACTCTCAC 5804
QY 1793 CTGCTGGGTTCAAGATTTTCTGCTGCTGAGCTCTCCAGTACTGGGAATACAGGACAG 1852
DB 5803 CTGCTGGGTTCAAGATTTTCTGCTGCTGAGCTCTCCAGTACTGGGAATACAGGACAG 5744
QY 1853 CCACCATGCCCAGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTT 1912
DB 5743 CCACCATGCCCAGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTT 5684
QY 1913 GGTGGCTGTGCAACTCTCTGACCGCAAGTATCCACCCTGCTGCTCCCAAGTCTGG 1972
DB 5683 GGTGGCTGTGCAACTCTCTGACCGCAAGTATCCACCCTGCTGCTCCCAAGTCTGG 5624
QY 1973 GATTACAGGCGTGAAGCAGCGGCGCCGCAAGGGGAACTCTGTGGAGAGACAA 2032
DB 5623 GATTACAGGCGTGAAGCAGCGGCGCCGCAAGGGGAACTCTGTGGAGAGACAA 5564
QY 2033 GGGGCTCACATCTCCCTCTGATTTCCCATGACATTCCTATCTCTCCCATCTAGC 2092
DB 5563 GGGGCTCACATCTCCCTCTGATTTCCCATGACATTCCTATCTCTCCCATCTAGC 5504
QY 2093 CAGGAATCTATTGTGTTTTTCTTCTGCAATTTTACTATGATGTATGTATGTATGTATGTAT 2152
DB 5503 CAGGAATCTATTGTGTTTTTCTTCTGCAATTTTACTATGATGTATGTATGTATGTATGTAT 5444
QY 2153 CCAACCCCGCCAGTGGGGGGTGGAGAGGGGTGCAAGCCCTGCTGCTCACTTTTCTTA 2212
DB 5443 CCAACCCCGCCAGTGGGGGGTGGAGAGGGGTGCAAGCCCTGCTGCTCACTTTTCTTA 5384
QY 2213 CCTTGAAGTGTATGTATTAATCACTTCTGTTTGTTCAGTTTTC 2260
DB 5383 CCTTGAAGTGTATGTATTAATCACTTCTGTTTGTTCAGTTTTC 5336

RESULT 12

US-09-764-904-90/c
Sequence 90, Application US/09764904
Patent No. US20020173454A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA122
CURRENT APPLICATION NUMBER: US/09/764,904
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 137
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 90

LENGTH: 32174
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (29356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29357)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29358)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29360)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29361)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29362)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29363)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29364)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29365)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29366)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29367)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29368)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29369)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29370)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29371)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29372)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29373)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29374)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29375)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29376)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29377)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29378)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (29379) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29380) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29381) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29382) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29383) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29384) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29385) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29386) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29387) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29388) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29389) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29390) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29391) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29392) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29393) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29394) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29395) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29396) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29397) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29398) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29399) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29400) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29401) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29402) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE

LOCATION: (29403) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29404) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29405) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29406) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29407) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29408) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29409) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29410) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29411) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29412) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29413) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29414) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29415) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29416) OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE

Query Match 31.1%; Score 708; DB 10; Length 32174;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GGTGGCTGTGCGGAGCTGTGAGAACTACTGTCCTTTTCACAGCACTCCTTG 1612
DB 6043 GGTGGCTGTGCGGAGCTGTGAGAACTACTGTCCTTTTCACAGCACTCCTTG 5984
QY 1613 CCCGAGCAGAGATGAGAAAGCCAGGAGGTGGAAGATCGATGCTTCCAGCTGTGCT 1672
DB 5983 CCCGAGCAGAGATGAGAAAGCCAGGAGGTGGAAGATCGATGCTTCCAGCTGTGCT 5924
QY 1673 CTGCTGCCAGCAGATCTTATTTGGGCAAGGGGAACTTTTGGAGAAAGGCT 1732
DB 5923 CTGCTGCCAGCAGATCTTATTTGGGCAAGGGGAACTTTTGGAGAAAGGCT 5864
QY 1733 CTGCTTTTGCACCCAGCTGATGAGTGGGGGATCTCAGCTCAGCCAACTTCCAC 1792
DB 5863 CTGCTTTTGCACCCAGCTGATGAGTGGGGGATCTCAGCTCAGCCAACTTCCAC 5804
QY 1793 CTCTGGGTCAAGATTTTCTGCTCAGCTTCCAAATGAGTGAATACAGGCACG 1852
DB 5803 CTCTGGGTCAAGATTTTCTGCTCAGCTTCCAAATGAGTGAATACAGGCACG 5744
QY 1853 CCACATGCCAGCTAATTTTGTATTTTCAATGAGAAAGGGATTTTCAACAGCTTGGCCA 1912
DB 5743 CCACATGCCAGCTAATTTTGTATTTTCAATGAGAAAGGGATTTTCAACAGCTTGGCCA 5684
QY 1913 GGCTGCTCGAATCTCTGACGCAAGTATCCACCCGCTCGCTCCCAAGTGTG 1972

Db 5683 GGGTGTCTCGAACTCTCTGACCCGCAAGTATCCACCCGCTCCGCTCCCAAGTGTGG 5624
 Qy 1973 GATTACAGGGGTGAGGACACCGTGGCCGCAAGGGGAACTCTGTGGAGAGACCA 2032
 Db 5623 GATTACAGGGGTGAGGACACCGTGGCCGCAAGGGGAACTCTGTGGAGAGACCA 5564
 Qy 2033 GGGGCTCACATCTCCCTCTGATTCGCCATGACATTCCTTATCTCCCATCTAGC 2092
 Db 5563 GGGGCTCACATCTCCCTCTGATTCGCCATGACATTCCTTATCTCCCATCTAGC 5504
 Qy 2093 CAGGATCTATGTGTGTCTCTCTGCAATTTACTATGATGTGTATGTGCGCTACCA 2152
 Db 5503 CAGGATCTATGTGTGTCTCTCTGCAATTTACTATGATGTGTATGTGCGCTACCA 5444
 Qy 2153 CCACCCCGCCATGGGGGGGTGAGAGGGGTGCAAGCCCTGCTGCTCACTTTTCA 2212
 Db 5443 CCACCCCGCCATGGGGGGGTGAGAGGGGTGCAAGCCCTGCTGCTCACTTTTCA 5384
 Qy 2213 CCTTGGAAGTATGATTAATAATCACTTCTGTTGTGAGTTTCA 2260
 Db 5383 CCTTGGAAGTATGATTAATAATCACTTCTGTTGTGAGTTTCA 5336

RESULT 13
 US-09-764-891-6480/c
 Sequence 6480, Application US/09764891
 Publication No. US20030077808A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PC006
 CURRENT APPLICATION NUMBER: US/09/764,891
 CURRENT FILING DATE: 2001-01-17
 Prior application data removed - consult PALM or file wrapper
 NUMBER OF SEQ ID NOS: 10231
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 6480
 LENGTH: 32174
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (29356)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29357)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29358)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29359)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29360)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29361)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29362)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29363)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29364)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29365)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29366)

OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29367)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29368)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29369)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29370)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29371)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29372)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29373)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29374)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29375)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29376)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29377)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29378)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29379)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29380)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29381)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29382)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29383)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29384)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29385)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29386)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29387)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29388)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29389)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (29390)
 OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (29391) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29392) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29393) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29394) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29395) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29396) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29397) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29398) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29399) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29400) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29401) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29402) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29403) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29404) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29405) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29406) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29407) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29408) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29409) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29410) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29411) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29412) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29413) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29414) n equals a,t,g, or c
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE

LOCATION: (29415)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29416)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE

Query Match 31.1%; Score 708; DB 11; Length 32174;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1553 GGTGGCTGTGGGGGAGCTGTGCGAAGAACTACGCTCTCCCTTTTACAGCACTCTTTG 1612
6043 GGTGGCTGTGGGGGAGCTGTGCGAAGAACTACGCTCTCCCTTTTACAGCACTCTTTG 5984
1613 CCCAGAGCAGAGAAATGAGAAAGCCAGGAGGTGAGATGATGCTTCACGCTGTGCTT 1672
5983 CCCAGAGCAGAGAAATGAGAAAGCCAGGAGGTGAGATGATGCTTCACGCTGTGCTT 5924
1673 CTGCTGCGCAGCAAGCTTCATTTGGGGCCAAAGGGGAACTTTTGTGAGAGGCGT 1732
5923 CTGCTGCGCAGCAAGCTTCATTTGGGGCCAAAGGGGAACTTTTGTGAGAGGCGT 5864
1733 CTGCTGCGCAGCAAGCTTCATTTGGGGCCAAAGGGGAACTTTTGTGAGAGGCGT 1792
5863 CTGCTGCGCAGCAAGCTTCATTTGGGGCCAAAGGGGAACTTTTGTGAGAGGCGT 5804
1793 CTGCTGCGCAGCAAGCTTCATTTGGGGCCAAAGGGGAACTTTTGTGAGAGGCGT 1852
5803 CTGCTGCGCAGCAAGCTTCATTTGGGGCCAAAGGGGAACTTTTGTGAGAGGCGT 5744
1853 CCACATGCGCCAGCTTAATTTTGTATTTAGTAGAAAGGATTTACACAGCTGGCCA 1912
5743 CCACATGCGCCAGCTTAATTTTGTATTTAGTAGAAAGGATTTACACAGCTGGCCA 5684
1913 GGCTGCTCGAATCTCTGACGCAAGTATCCACCGCTCCGCTCCCAAGTGTGG 1972
5683 GGCTGCTCGAATCTCTGACGCAAGTATCCACCGCTCCGCTCCCAAGTGTGG 5624
1973 GATTACAGCGGTGAGGACCGCTGCGGCGCCAAAGGGGAACTTTGTGGAGAGCAGA 2032
5623 GATTACAGCGGTGAGGACCGCTGCGGCGCCAAAGGGGAACTTTGTGGAGAGCAGA 5564
2033 GGGGCTCACATCTCCCTCTGATTCGCCCATGACATTCCTATCTCCCATATGAC 2092
5563 GGGGCTCACATCTCCCTCTGATTCGCCCATGACATTCCTATCTCCCATATGAC 5504
2093 CAGGATCTATTTGTTTTCTTCTGCAATTTATGATGATGATGATGCGGTACCA 2152
5503 CAGGATCTATTTGTTTTCTTCTGCAATTTATGATGATGATGATGCGGTACCA 5444
2153 CCACCCCCCAGTGGGGGGGTGAGAGGGGTGCAAGCCCTGCTCTCACTTTTCTA 2212
5443 CCACCCCCCAGTGGGGGGGTGAGAGGGGTGCAAGCCCTGCTCTCACTTTTCTA 5384
2213 CCTGGAAGTATGATTAATCACTCTGTTGTTGAGTTTCA 2260
5383 CCTGGAAGTATGATTAATCACTCTGTTGTTGAGTTTCA 5336

RESULT 14
US-09-764-891-10135
Sequence 10135, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
PRIORITY DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO	10135
LENGTH	32174
TYPE	DNA
ORGANISM	Homo sapiens
FEATURE:	
NAME/KEY:	SITE
LOCATION:	(1295)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1296)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1297)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1298)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1299)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1300)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1301)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1302)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1303)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1304)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1305)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1306)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1307)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1308)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1309)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1310)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1311)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1312)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1313)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1314)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1315)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1316)
OTHER INFORMATION:	n equals a,t,g, or c
NAME/KEY:	SITE
LOCATION:	(1317)
OTHER INFORMATION:	n equals a,t,g, or c

```
NAME/KEY: SITE
LOCATION: (1342)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1344)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1346)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1347)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1349)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1350)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1351)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1352)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1353)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1354)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1355)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
```

Query Match 31.1%; Score 708; DB 11; Length 32174;

Best Local Similarity 100.0%; Pred.No. 0;

Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1553 GGTGGCTGTGGGAGCTGTGGAGAACTAGTCTCCCTTTTCAAGCACTCTTTG 1612
26132 GGTGGCTGTGGGAGCTGTGGAGAACTAGTCTCCCTTTTCAAGCACTCTTTG 26191
1613 CCCCAGAGCAGAGATGAGAAAGCCAGAGAGGTGAGATGATGCTTCCAGCTGTGCT 1672
26192 CCCCAGAGCAGAGATGAGAAAGCCAGAGAGGTGAGATGATGCTTCCAGCTGTGCT 26251
1673 CTGCTGCCAGCAGATCTTCACTTTGGGGCCAAAGGGGAACTTTTGTGAGAGGGCT 1732
26252 CTGCTGCCAGCAGATCTTCACTTTGGGGCCAAAGGGGAACTTTTGTGAGAGGGCT 26311
1733 CTGCTTTGTCAAGCAGGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1792
26312 CTGCTTTGTCAAGCAGGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 26371
1793 CTCTGGGTTCAAGTGAATTTTCTGCTCAGCTTCCCAAGTGTGGGATTCAGAGCGAC 1852
26372 CTCTGGGTTCAAGTGAATTTTCTGCTCAGCTTCCCAAGTGTGGGATTCAGAGCGAC 26431
1853 CCACCATGCCAGCTAATTTTGTATTTTCACTAGAGAGGATTTCAACGATGGGCA 1912
26432 CCACCATGCCAGCTAATTTTGTATTTTCACTAGAGAGGATTTCAACGATGGGCA 26491
1913 GGCTGGTTCGAATCTCTGACGCAAGTATCAACCCGCTCCGCTCCCAAGTGTGG 1972
```

```
26492 GGTGGTTCGAATCTCTGACGCAAGTATCAACCCGCTCCGCTCCCAAGTGTGG 26551
1973 GATTACAGCGTGAAGCAGCAGTCCCGGCCAAAGGGGAACTTTGTGGAGAGCAGA 2032
26552 GATTACAGCGTGAAGCAGCAGTCCCGGCCAAAGGGGAACTTTGTGGAGAGCAGA 26611
2033 GGGGCTCAATCTCCCTCTGATTTCCCATGCACTTGCCTTATCTTCCCATATAGC 2092
26612 GGGGCTCAATCTCCCTCTGATTTCCCATGCACTTGCCTTATCTTCCCATATAGC 26671
2093 CAGAACTATTGTGTTTCTTCTGCAATTTACTATGATTTGATGTCGCGTACCA 2152
26672 CAGAACTATTGTGTTTCTTCTGCAATTTACTATGATTTGATGTCGCGTACCA 26731
2153 CCACCCCCCATGAGGGGGGTGAGAGGGGTGCAAGCCCTGCTCCACTTTTCTA 2212
26732 CCACCCCCCATGAGGGGGGTGAGAGGGGTGCAAGCCCTGCTCCACTTTTCTA 26791
2213 CCTTGAATCTGATTTGATTAATCACTTCTGTTGTGATTTTCA 2260
26792 CCTTGAATCTGATTTGATTAATCACTTCTGTTGTGATTTTCA 26839
```

RESULT 15

```
US-09-764-891-10179/c
Sequence 10179, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10179
LENGTH: 32174
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (29356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29357)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29358)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29360)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29361)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29362)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29363)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29364)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (29365)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
```

OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29391)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29392)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29393)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29394)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29395)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29396)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29397)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29398)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29399)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29400)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29401)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29402)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29403)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29404)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29405)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29406)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29407)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29408)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29409)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29410)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29411)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29412)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29413)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29414)
OTHER INFORMATION: n equals a,t,g, or c	NAME/KEY: SITE	LOCATION: (29415)

```

; NAME/KEY: SITE
; LOCATION: (29415)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (29416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

```

```

Query Match      31.1%; Score 708; DB 11; Length 32174;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1553 GGTGGCTGTGGGAGTGTGCGAAGAACTGCTCCCTTTTCACAGCACTCCTTTG 1612
DB 6043 GGTGGCTGTGGGAGTGTGCGAAGAACTGCTCCCTTTTCACAGCACTCCTTTG 5984
QY 1613 CCCAGAGCAGAGATGAGAAAGCCAGAGAGTGAATGATGCTTCCAGCTGTGCT 1672
DB 5983 CCCAGAGCAGAGATGAGAAAGCCAGAGAGTGAATGATGCTTCCAGCTGTGCT 5924
QY 1673 CTGCTGCAGCAGCAAGTCTTATTTGGGCGCAAAAGGGGAACTTTTGGAGAGCGT 1732
DB 5923 CTGCTGCAGCAGCAAGTCTTATTTGGGCGCAAAAGGGGAACTTTTGGAGAGCGT 5864
QY 1733 CTTGCTTTGTCACTCCAGCGCTGGAATGCAGTGGCGGATCTGAGCTCAACCGCACTCCAC 1792
DB 5863 CTTGCTTTGTCACTCCAGCGCTGGAATGCAGTGGCGGATCTGAGCTCAACCGCACTCCAC 5804
QY 1793 CTCCTGGGTTCAAGTATTTTCCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG 1852
DB 5803 CTCCTGGGTTCAAGTATTTTCCTGCTCAGCTCCCAAGTAGCTGGGAATACAGGACG 5744
QY 1853 CCACCATGCCCAGCAATTTTGTATTTTCAGTAGAAAAGGGATTTCCACAGTTGGCCA 1912
DB 5743 CCACCATGCCCAGCAATTTTGTATTTTCAGTAGAAAAGGGATTTCCACAGTTGGCCA 5684
QY 1913 GGTGCTGTGGAATCTCTGACCGCAAGTATCCACCGGCTCCGCTCCCAAGTGTGG 1972
DB 5683 GGTGCTGTGGAATCTCTGACCGCAAGTATCCACCGGCTCCGCTCCCAAGTGTGG 5624
QY 1973 GATTACAGGCGTAGGCCACGCTGCCGCGCAAAAGGGGAACTTTGTGGAGAGCAGA 2032
DB 5623 GATTACAGGCGTAGGCCACGCTGCCGCGCAAAAGGGGAACTTTGTGGAGAGCAGA 5564
QY 2033 GGGGCTCAATCTCCCTCTGATTCGCCATGACATTCCTTATCTCTCCCATCTAGC 2092
DB 5563 GGGGCTCAATCTCCCTCTGATTCGCCATGACATTCCTTATCTCTCCCATCTAGC 5504
QY 2093 CAGGAATCTATTTGTTTTTCTTCCCAATTACTATGATGTGTATGCGCGCTACCA 2152
DB 5503 CAGGAATCTATTTGTTTTTCTTCCCAATTACTATGATGTGTATGCGCGCTACCA 5444
QY 2153 CCACCCCCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTCCTCACTTTTCTA 2212
DB 5443 CCACCCCCCCCCATGGGGGGGTGAGAGGGGTGCAAGGCCCTGCTCCTCACTTTTCTA 5384
QY 2213 CTTTGAAGTGTATGATAAATCACTTCTGTTTGTCAAGTTTCA 2260
DB 5383 CTTTGAAGTGTATGATAAATCACTTCTGTTTGTCAAGTTTCA 5336

```

Search completed: November 22, 2003, 14:21:35
Job time : 707 sec

APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Kijavlin, Ivar J.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2730P1C13
 CURRENT APPLICATION NUMBER: US/09/996,243
 CURRENT FILING DATE: 2001-11-14
 PRIOR APPLICATION NUMBER: 60/049787
 PRIOR FILING DATE: 1997-06-16
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/075945
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09

PRIOR APPLICATION NUMBER: 60/088734
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088742
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 2.2%; Score 49; DB 4; Length 907;
Best Local Similarity 100.0%; Pred. No. 2,8e-12;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTCTGGATTACAGCGTGAGCCACCGTCCCGGCCCA 2004
DB 668 GCCTCCCAAGTCTGGATTACAGCGTGAGCCACCGTCCCGGCCCA 620

RESULT 3

US-09-804-471A-3
Sequence 3, Application US/09804471A
Patent No. 6479269
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001164
CURRENT APPLICATION NUMBER: US/09/804,471A
CURRENT FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 174493
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(174493)
OTHER INFORMATION: n = A,T,C or G

US-09-804-471A-3

Query Match 2.2%; Score 49; DB 4; Length 174493;
Best Local Similarity 100.0%; Pred. No. 2,4e-12;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTCTGGATTACAGCGTGAGCCACCGTCCCGGCCCA 2004
DB 23465 GCCTCCCAAGTCTGGATTACAGCGTGAGCCACCGTCCCGGCCCA 23513

RESULT 4

US-09-305-384-7/C
Sequence 7, Application US/09305384
Patent No. 6242218
GENERAL INFORMATION:
APPLICANT: Treco, Douglas A.
APPLICANT: Hearlein, Michael W.
APPLICANT: Selden, Richard F.
TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
FILE REFERENCE: 07236/017001
CURRENT APPLICATION NUMBER: US/09/305,384
CURRENT FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,649
PRIOR FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 1252
TYPE: DNA
ORGANISM: Homo sapiens
US-09-305-384-7

Query Match 2.1%; Score 47; DB 3; Length 1252;
Best Local Similarity 100.0%; Pred. No. 2,1e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTCTGGATTACAGCGTGAGCCACCGTCCCGGCCC 2002
DB 524 GCCTCCCAAGTCTGGATTACAGCGTGAGCCACCGTCCCGGCCC 478

RESULT 5

US-09-702-705-321
Sequence 321, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 321
LENGTH: 2280
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-321

Query Match 2.1%; Score 47; DB 4; Length 2280;
Best Local Similarity 100.0%; Pred. No. 2,1e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 2002
 DB 1113 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 1159

RESULT 6

US-09-736-457-321
 ; Sequence 321, Application US/09736457
 ; Patent No. 6509448
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darlick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Pan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.478C15
 ; CURRENT APPLICATION NUMBER: US/09/736,457
 ; NUMBER OF SEQ ID NOS: 1864
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 321
 ; LENGTH: 2280
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-736-457-321

Query Match 2.1%; Score 47; DB 4; Length 2280;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 2002
 DB 1113 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 1159

RESULT 7

US-10-045-428A-9
 ; Sequence 9, Application US/10045428A
 ; Patent No. 6500943
 ; GENERAL INFORMATION:
 ; APPLICANT: Mano, Hiroyuki
 ; APPLICANT: Sakata, Tsunaki
 ; APPLICANT: Hasegawa, Mamoru
 ; APPLICANT: Tabata, Toshiaki
 ; TITLE OF INVENTION: Promoter
 ; FILE REFERENCE: 50026/011003
 ; CURRENT APPLICATION NUMBER: US/10/045,428A
 ; CURRENT FILING DATE: 2002-04-15
 ; PRIOR APPLICATION NUMBER: 09/735,103
 ; PRIOR FILING DATE: 2000-12-12
 ; PRIOR APPLICATION NUMBER: 09/142,529
 ; PRIOR FILING DATE: 1999-08-12
 ; PRIOR APPLICATION NUMBER: PCT/JP97/00741
 ; PRIOR FILING DATE: 1997-03-10
 ; PRIOR APPLICATION NUMBER: JP 8/54294
 ; PRIOR FILING DATE: 1996-03-12
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 3129
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-045-428A-9

Query Match 2.1%; Score 47; DB 4; Length 3129;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1956 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 2002
 DB 410 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 456

RESULT 8

US-09-305-384-5/c
 ; Sequence 5, Application US/09305384
 ; Patent No. 6242218
 ; GENERAL INFORMATION:
 ; APPLICANT: Treco, Douglas A.
 ; APPLICANT: Heutlein, Michael W.
 ; APPLICANT: Seiden, Richard F.
 ; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
 ; FILE REFERENCE: 07236/017001
 ; CURRENT APPLICATION NUMBER: US/09/305,384
 ; CURRENT FILING DATE: 1999-05-05
 ; EARLIER APPLICATION NUMBER: US 60/084,649
 ; EARLIER FILING DATE: 1998-05-07
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 6235
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-305-384-5

Query Match 2.1%; Score 47; DB 3; Length 6235;
 Best Local Similarity 100.0%; Pred. No. 2e-11;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 2002
 DB 5251 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 5205

RESULT 9

US-09-525-160B-6/c
 ; Sequence 6, Application US/09525160B
 ; Patent No. 6569681
 ; GENERAL INFORMATION:
 ; APPLICANT: Ivanov, Eugene
 ; TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION
 ; FILE REFERENCE: 10278/016001
 ; CURRENT APPLICATION NUMBER: US/09/525,160B
 ; CURRENT FILING DATE: 2000-03-14
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 6235
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-525-160B-6

Query Match 2.1%; Score 47; DB 4; Length 6235;
 Best Local Similarity 100.0%; Pred. No. 2e-11;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 2002
 DB 5251 GCCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGTCCCGGCC 5205

RESULT 10

US-09-305-384-1/c
 ; Sequence 1, Application US/09305384
 ; Patent No. 6242218
 ; GENERAL INFORMATION:
 ; APPLICANT: Treco, Douglas A.
 ; APPLICANT: Heutlein, Michael W.
 ; APPLICANT: Seiden, Richard F.

TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY

FILE REFERENCE: 07236/017001

CURRENT APPLICATION NUMBER: US/09/305,384

CURRENT FILING DATE: 1999-05-05

EARLIER APPLICATION NUMBER: US 60/084,649

EARLIER FILING DATE: 1998-05-07

NUMBER OF SEQ ID NOS: 8

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 6679

TYPE: DNA

ORGANISM: Homo sapiens

US-09-305-384-1

Query Match 2.1%; Score 47; DB 3; Length 6679;

Best Local Similarity 100.0%; Pred. No. 2e-11;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTCTGGATTACAGCGCGTGCACCGTCCCGGCC 2002

DB 5270 GCCTCCCAAGTCTGGATTACAGCGCGTGCACCGTCCCGGCC 5224

RESULT 11

US-09-525-160B-5/C

Sequence 5, Application US/09525160B

Patent No. 6569681

GENERAL INFORMATION:

APPLICANT: Ivanov, Evgenii

TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION

FILE REFERENCE: 10278/016001

CURRENT APPLICATION NUMBER: US/09/525,160B

CURRENT FILING DATE: 2000-03-14

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 6679

TYPE: DNA

ORGANISM: Homo sapiens

US-09-525-160B-5

Query Match 2.1%; Score 47; DB 4; Length 6679;

Best Local Similarity 100.0%; Pred. No. 2e-11;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTCTGGATTACAGCGCGTGCACCGTCCCGGCC 2002

DB 5270 GCCTCCCAAGTCTGGATTACAGCGCGTGCACCGTCCCGGCC 5224

RESULT 12

US-08-257-963B-10/C

Sequence 10, Application US/08257963B

Patent No. 5840686

GENERAL INFORMATION:

APPLICANT: Chader, Gerald J.; Becerra, S.

APPLICANT: Patricia; Schwartz, Joan P.;

APPLICANT: Taniwaki, Takayuki

TITLE OF INVENTION: PIGMENT EPITHELIUM

TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL

TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan

STREET: 345 Park Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/257,963B

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/952,796

FILING DATE: 24-SEPT-1992

ATTORNEY/AGENT INFORMATION:

NAME: DOROTHY R. AUTH

REGISTRATION NUMBER: 36434

REFERENCE/DOCKET NUMBER: 20264126US1

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 7210 Base Pairs

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Unknown

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE:

ORGANISM: Human

IMMEDIATE SOURCE:

LIBRARY: DASH II

FEATURE:

NAME/KEY: JT106

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: 7.2 kb No. 5840686 1 fragments

OTHER INFORMATION: Derived from human placental genomic DNA

US-08-257-963B-10

Query Match 2.1%; Score 47; DB 2; Length 7210;

Best Local Similarity 100.0%; Pred. No. 2e-11;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTCTGGATTACAGCGCGTGCACCGTCCCGGCC 2002

DB 5615 GCCTCCCAAGTCTGGATTACAGCGCGTGCACCGTCCCGGCC 5569

RESULT 13

US-08-367-841A-10/C

Sequence 10, Application US/08367841A

Patent No. 6319687

GENERAL INFORMATION:

APPLICANT: Chader, Gerald J.; Rodriguez,

APPLICANT: Ignacio R.; Mazuruk, Krzysztof;

APPLICANT: Tombran-Tink, Joyce

TITLE OF INVENTION: PIGMENT EPITHELIUM

TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC

TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan

STREET: 345 Park Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/367,841A

FILING DATE: 30-DEC-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/257,963
FILING DATE: 07-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/952,796
FILING DATE: 24-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36434
REFERENCE/DOCKET NUMBER: 20264126US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 7210 Base Pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Human
IMMEDIATE SOURCE:
LIBRARY: DASH II
FEATURE:
NAME/KEY: JTEA
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: 7.0 kb No. 6319687 1-No. 6319687
OTHER INFORMATION: fragment; derived from human placental
OTHER INFORMATION: genomic DNA; also referred to as JTI06
US-08-367-841A-10

Query Match 2.1%; Score 47; DB 4; Length 7210;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1956 GCCTCCCAAGTCTGGATTACAGCGGTGACCCGTCGCCGCC 2002
DB 5615 GCCTCCCAAGTCTGGATTACAGCGGTGACCCGTCGCCGCC 5569

RESULT 14
PCT-US95-07201-10/c
Sequence 10, Application PC/TUS9507201
GENERAL INFORMATION:
APPLICANT: Chader, Gerald J.; Becerra, Sofia
APPLICANT: Patricia; Schwartz, Joan P.;
APPLICANT: Tanigaki, Takayuki
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07201
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/367,841
FILING DATE: 30-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,963

FILING DATE: 07-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/952,796
FILING DATE: 24-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36434
REFERENCE/DOCKET NUMBER: 20264126PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 7210 Base Pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Human
IMMEDIATE SOURCE:
LIBRARY: DASH II
FEATURE:
NAME/KEY: JTEA
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: 7.0 kb Not 1-Not
OTHER INFORMATION: fragment; derived from human placental
OTHER INFORMATION: genomic DNA; also referred to as JTI06
PCT-US95-07201-10

Query Match 2.1%; Score 47; DB 5; Length 7210;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1956 GCCTCCCAAGTCTGGATTACAGCGGTGACCCGTCGCCGCC 2002
DB 5615 GCCTCCCAAGTCTGGATTACAGCGGTGACCCGTCGCCGCC 5569

RESULT 15
US-08-652-265-1/c
Sequence 1, Application US/08652265
Patent No. 6025130
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Dzayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Guirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,265
FILING DATE: 23-MAY-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,723

```

REFERENCE/DOCKET NUMBER: 17957-000500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /note= "No. 6025130mal or wild-type (unaffected)"
OTHER INFORMATION: Hereditary Hemochromatosis (HH) gene
FEATURE:
NAME/KEY:
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: normal or wild-type (unaffected) allele"
FEATURE:
NAME/KEY:
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: normal or wild-type (unaffected) genomic
OTHER INFORMATION: sequence surrounding variant for 24d2 (C)
FEATURE:
NAME/KEY:
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: normal or wild-type (unaffected) genomic
OTHER INFORMATION: sequence surrounding variant for 24d1 (G)
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "C")
OTHER INFORMATION: /phenotype= "normal or wild-type
OTHER INFORMATION: (unaffected)"
FEATURE:
NAME/KEY:
LOCATION:
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(3878, "A")
OTHER INFORMATION: /phenotype= "normal or wild-type
OTHER INFORMATION: (unaffected)"
FEATURE:
NAME/KEY:
LOCATION:
OTHER INFORMATION: /label= 24d7
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "G")
OTHER INFORMATION: /phenotype= "normal or wild-type
OTHER INFORMATION: (unaffected)"
OTHER INFORMATION: /label= 24d1
US-08-652-265-1

Query Match 2.1%; Score 47; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1956 GCCTCCCAAGTGCCTGATTACAGCGGTGAGCCACCGTCCCGGCC 2002
DB 8175 GCCTCCCAAGTGCCTGATTACAGCGGTGAGCCACCGTCCCGGCC 8129

```

Search completed: November 22, 2003, 14:01:17
 Job time: 137 secs

THIS PAGE BLANK (USPTO)